BEHAVIORAL AND BRAIN SCIENCES (2001) **24,** 849–937
*Printed in the United States of America*

# The Theory of Event Coding (TEC): A framework for perception and action planning

**Bernhard Hommel,[a,b] Jochen Müsseler,[b] Gisa Aschersleben,[b]** and **Wolfgang Prinz[b]**
[a]*Section of Experimental and Theoretical Psychology, University of Leiden, 2300 RB Leiden, The Netherlands;* [b]*Max Planck Institute for Psychological Research, D-80799 Munich, Germany*
**{muesseler;aschersleben;prinz}@mpipf-muenchen.mpg.de**
**www.mpipf-muenchen.mpg.de/~prinz** **hommel@fsw.leidenuniv.nl**

**Abstract:** Traditional approaches to human information processing tend to deal with perception and action planning in isolation, so that an adequate account of the perception-action interface is still missing. On the perceptual side, the dominant cognitive view largely underestimates, and thus fails to account for, the impact of action-related processes on both the processing of perceptual information and on perceptual learning. On the action side, most approaches conceive of action planning as a mere continuation of stimulus processing, thus failing to account for the goal-directedness of even the simplest reaction in an experimental task. We propose a new framework for a more adequate theoretical treatment of perception and action planning, in which perceptual contents and action plans are coded in a common representational medium by feature codes with distal reference. Perceived events (perceptions) and to-be-produced events (actions) are equally represented by integrated, task-tuned networks of feature codes – cognitive structures we call event codes. We give an overview of evidence from a wide variety of empirical domains, such as spatial stimulus-response compatibility, sensorimotor synchronization, and ideomotor action, showing that our main assumptions are well supported by the data.

**Keywords:** action planning; binding; common coding; event coding; feature integration; perception; perception-action interface

## 1. Introduction and overview

We propose a new theoretical framework for the cognitive underpinnings of perception and action planning, the Theory of Event Coding (TEC). Basically, the theory holds that cognitive representations of events (i.e., of any to-be-perceived or to-be-generated incident in the distal environment) subserve not only representational functions (e.g., for perception, imagery, memory, reasoning) but action-related functions as well (e.g., for action planning and initiation). According to TEC, the core structure of the functional architecture supporting perception and action planning is formed by a common representational domain for perceived events (perception) and intended or to-be-generated events (action).

In a nutshell, we believe that it makes sense to assume that the stimulus representations underlying perception and the action representations underlying action planning are coded and stored not separately, but together in a common representational medium. This implies that stimulus and response codes are not entities of a completely different kind, but only refer to, and thus represent, different events in a particular task and context. Thus, it would be misleading to speak of stimulus codes and response or action codes unless one wishes to refer to the roles played by a code or the event it represents (see sect. 3.2.4). Irrespective of this role, though, cognitive codes are always event codes – codes of perceived or (to-be-)produced events.

Though TEC is meant to provide a new framework for perception and action planning, its scope is limited in the following sense: As regards perception, its focus is on "late" cognitive products of perceptual processing that stand for, or represent, certain features of actual events in the environment. TEC does not consider the complex machinery of the "early" sensory processes that lead to them. Conversely, as regards action, the focus is on "early" cognitive antecedents of action that stand for, or represent, certain features of events that are to be generated in the environment (= actions). TEC does not consider the complex machinery of the "late" motor processes that subserve their realization (i.e., the control and coordination of movements). Thus, TEC is meant to provide a framework for understanding linkages between (late) perception and (early) action, or action planning. Therefore, we do not claim that TEC covers all kinds of interactions between perception and action exhaustively. The same applies to the representationalist approach inherent in TEC. Though we do believe that the representationalist stance that we adopt forms an appropriate metatheoretical framework for our theory, we do not want to imply, or suggest, that it is necessary or appropriate for understanding other kinds of interactions between perception and action.

As we will point out below, TEC differs from other, in part related approaches to perception and action planning. In contrast to the classical information processing view it does not see perception and action planning as different,

© 2001 Cambridge University Press 0140-525X/01 $12.50

https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

functionally separable stages but as intimately related, sometimes even indistinguishable processes producing a wide variety of interactions. This perspective TEC shares with the ecological approach, from which it differs, however, by its emphasis on representational issues – which are anathema to ecological approaches by definition. Finally, TEC does not deal with the question of how consciousness and externally guided pointing and grasping movements are related, a topic gaining its attraction from demonstrations that both can be dissociated under specific conditions (for reviews, see Milner & Goodale 1995; Rossetti & Revonsuo 2000). Although these demonstrations are interesting and challenging, we do not see how they could provide a basis for the general theory of intentional action we aim at.

Our argument for TEC will take three major steps. In the first step we review classical approaches to perception and action planning, with special emphasis on their mutual linkages. As regards perception, this review will show that cognitive approaches to perception that are based on linear stage theory and its derivatives do not give an adequate account of the interface between perception and action planning. Unlike approaches in the domain of spatial orientation they neither have a way to accommodate the interaction between perception and action planning in a satisfactory way nor to account for the impact of action-related knowledge on perception. As regards action and action control, our review will also show that classical approaches are insufficient on both theoretical and empirical grounds. They neither provide an adequate role for action goals nor can they account for evidence witnessing the operation of similarity between perception and action planning. Based on these two reviews we will conclude that a new framework is needed for taking care of the cognitive underpinnings of the mutual interaction between perception and action planning.

Our review of shortcomings of classical theories will help us to put constraints on a new framework claiming to account for these issues in a more adequate way. As a result, TEC is presented in the second step. This theory proposes as its core contention that codes of perceived events and planned actions share a common representational domain, to the effect that perceptual codes and action codes may prime each other on the basis of their overlap in this domain. The structural view underlying this notion regards event codes as assemblies of feature codes, based on temporary integration in a given task context.

As we will show, TEC is not entirely new. It rather takes up elements from old and forgotten theories and (re)combines them in a novel way. This is particularly true of the ideomotor principle (cf. Greenwald 1970) and of motor theories of cognition that were so broadly discussed a century ago (cf. Scheerer 1984) and then fell into oblivion because the theoretical *zeitgeist* of the new century was input-centered throughout and did not much care about action per se. Precursor versions of TEC and treatments of some aspects of its functional logic have been published before (cf. Aschersleben & Prinz 1995; 1997; Hommel 1997; 1998a; Müsseler 1995; 1999; Prinz 1984b; 1987; 1990; 1997a; 1997b; Prinz et al. 1995). Its present version has emerged from our collaboration on a number of experimental projects and our efforts to understand the major theoretical implications of their results.

In the third step, we will then apply the framework of TEC to a number of pertinent experimental paradigms. As TEC is more a loose framework of theoretical principles rather than a strict, formal theory, there is no obvious way to test its validity directly, that is, by deriving predictions from it and then confronting those predictions with data. Instead, an additional step of translation will often be required that boils the principles of the general theory down to task-specific models. As a consequence the general theory can only be indirectly tested: It stands the test to the extent that task-specific models, which embody its functional logic, prove to be valid. Unlike specific models, which can be tested for empirical validity, the general theory behind them can only be evaluated in terms of the scope and the variety of successful models derived from it, that is, in terms of its heuristic and explanatory power. Thus, TEC is thought to serve a similar function as the linear-stage approach à la Donders (1868/1969) and Sternberg (1969), which also does not aim at explaining particular effects or phenomena, but rather provides a general framework and a theoretical language for developing more detailed and testable models and theories.

Most of the studies to which we will refer in the empiri-

Bernhard Hommel studied psychology and literature at the University of Bielefeld, where he also worked as a research assistant from 1987–1990 and completed his dissertation. He then moved to the Max Planck Institute for Psychological Research in Munich to work as a senior researcher. Since 1999 he has been a Full Professor (Chair of General Psychology) at the University of Leiden, Netherlands. He has published empirical and theoretical work on human attention, planning and control of action, and the relationship between perception and action.

Jochen Müsseler is senior scientist at the Max Planck Institute for Psychological Research (Department of Cognition and Action) in Munich. After graduating from the University of Bochum, he obtained his Ph.D. at the University of Bielefeld before moving to Munich. His research focuses on the interface of perception and action, the perception of space and time, and attentional mechanisms.

Gisa Aschersleben studied psychology at the Universities of Bielefeld and Braunschweig, Germany, and completed her Ph.D. at the University of Munich. She worked as a research assistant at the Technical University of Munich 1987–1991. In 1991 she joined the Max Planck Institute for Psychological Research in Munich where she holds the position of a senior scientist. In 2000 she became head of the research group Infant Cognition and Action. Her main research interests are in perception-action coupling, control of action, attention, intersensory integration, and early development of action control.

Wolfgang Prinz studied psychology, philosophy, and zoology at the University of Muenster, Germany. He took his Ph.D. in 1970 at the Department of Philosophy, Pedagogics and Psychology of the Ruhr-University, Bochum. He was a Professor of Psychology at the University of Bielefeld (1975–1990) and at the Ludwig-Maximilians-University Munich (1990–1998). Since 1990, he has been a Director at the Max Planck Institute for Psychological Research, Munich. He has published empirical and theoretical work on perception, action, conciousness, and attention as well as on the history of psychology.



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

cal section were originally inspired by TEC. At first glance, they seem to address a variety of disparate tasks like, for instance, sensorimotor synchronization, spatial S-R compatibility, identification and detection in backward masking, dual-task performance, ideomotor action, action perception, and action imitation. However, at a second, more theoretically inspired glance, they all address the same problem: mutual interactions between representations of perceived and produced events – and the structure of the functional architecture that takes care of generating, transforming, and storing these events.

We will conclude by pointing out some epistemological implications of our approach concerning the role of the interaction between perception and action planning for the construction of reality.

## 2. Perception and action planning

### 2.1. Views on perception

In this section we consider relationships between perception and action planning from the viewpoint of perception and perceptual theory. Perceptual views, though recognizing that perception and action planning may be highly interactive in many natural situations, tend to view perceptual processes in separation, that is, as more or less independent from processes of action planning and action control. This is not only true for the study of (presumably) early processes (e.g., figure-ground segregation, texture segmentation, or structural organization; cf. Beck 1982; Hochberg 1978; Metzger 1975; Palmer 1982; Pomerantz & Kubovy 1981; Restle 1982), but also for the study of (presumably) later processes in the chain of stimulus-induced operations that are responsible for identifying the stimulus or selecting appropriate responses. In the following sections we shall first examine how the linkage between perception and action is studied and conceptualized within the framework of information-processing approaches to perception (cf. Haber 1969; Massaro 1990; Posner 1978; Sanders 1980). These basically claim an information stream from perception to action with little contact between the two domains. We then survey views that stress the selective aspect of information processing: In order to perform goal-directed actions, this perceptual selection of certain aspects of environmental information is required, while other aspects are ignored or rejected. In that view, selection mechanisms take their way from action to perception. Finally, we consider perceptual mechanisms as part of a functional architecture whose outcome is adapted action. In this view, perceptual mechanisms are studied with respect to their contributions to people's or other animals' orientation in space and time. In this framework, perception and action thus have always been studied in common and in close correspondence.

**2.1.1. Information-processing views: From perception to action.** In his well-known paper "On the speed of mental processes," which appeared in 1868, the Dutch physiologist Frans Cornelis Donders established a nowadays widely accepted methodological strategy for decomposing the stream of processing between stimulus presentation and response generation into a number of stages. This method, that, according to Donders, allows us to compute the time components required by the "organ of perception" and the "organ of the will" for identifying the stimulus and selecting the response, respectively, is today considered one of the early important steps toward a natural science approach to mental phenomena – a major breakthrough in the early history of psychology. Donders' paper regained importance a hundred years later when the scientific community rediscovered its interest in the structure of the cognitive operations mediating between stimulus and response, an interest that originated in the 1950s and 1960s and continues till today. Since then, research in this field has seen an enormous growth of knowledge about the factors that determine stimulus processing and response generation, and we have gained a much broader understanding of the nature of the stages involved.

*2.1.1.1. Linear stage models.* Usually, stage-theoretical views encompass all cognitive processes that are settled along the information stream traveling between receptors and effectors, that is, from early perceptual processing of stimuli to the final execution of responses. Applying linear-stage logic allows us to establish a sequence of hypothetical processing stages and to assess their functional properties; it does so on the basis of experiments that study how task performance in speeded information transmission tasks depends on manipulations referring to features of the stimuli and/or demands of the task at hand (Massaro 1990; Sanders 1980; Sternberg 1969).

However, though a large number of cognitive psychologists subscribe to the information-processing approach, only few of them have taken interest in the total information stream between stimuli and responses. On the contrary, subdisciplines like perception, memory, problem solving, motor behavior, and so on, have preserved their relative autonomy. As far as perception is concerned, this autonomy emerges from the understanding that perceptual processing can be studied without taking later operations into account. For instance, it is generally believed (and it actually follows from the logic inherent in the approach) that stimulus-related processing can be studied completely independent from response-related processing. The reason is that the information stream in perception is commonly considered to be data-driven, in the sense of an uni-directional flow of information from peripheral to central stages.

*2.1.1.2. Perception, memory, and action.* The interaction between information derived from the stimulus and information stored in memory is a topic of central importance in information-processing models. In fact, it is this feature that qualifies them as cognitive models of perception. Such interaction is inherent in the notion of stimulus identification. It is usually believed that the stimulus gets identified by way of matching a stimulus representation against a set of memory representations. Classical studies have addressed the mechanisms underlying this interaction, raising issues like parallel versus serial, analytic versus holistic, or exhaustive versus self-terminating organization of matches in a number of tasks (Egeth 1966; Hick 1952; Neisser 1963; Nickerson 1972; Posner & Mitchell 1967; Sternberg 1967). Interestingly, these studies focus on processes rather than on contents and structure. They try to elucidate the nature of the operations underlying the interaction between stimulus and memory information, but they are much less explicit on the contents of the memory representations involved and the structural relations between them.

Such neglect of structure and content applies even more



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

when it comes to the memory requirements of the response-related operations subsequent to stimulus identification. Though the classical models do not explicitly specify how responses are coded and stored, they do presume, at least by implication, that response codes reside entirely separate from stimulus codes – stimulus codes standing for environmental events and response codes for body movements. For instance, in a choice-reaction time experiment, where participants press one of two keys (say, left vs. right-hand) in response to the color of a stimulus light (say, red vs. green), stimulus codes stand for colors and response codes for movements. Accordingly, a task like this can only be solved by means of a device that translates, as it were, colors into movements. The translation metaphor, which is widely used to characterize the nature of response selection (e.g., Massaro 1990; Sanders 1980; Welford 1968), thus stresses the incommensurability between stimulus and response.

*2.1.1.3. More elaborate flow-charts.* In the 1960s and 1970s, the linear stage model was taken literally. Stages were arranged in a row, and a given stage would not begin to work before the preceding stage had done its job (Donders 1868/1969; Sanders 1980; Sternberg 1969; Theios 1975). For example, Sanders (1983) proposed a linear stage model with four stages: stimulus preprocessing, feature extraction, response choice, and motor adjustment. Obviously, the first two stages are concerned with perception and the last two with response-related processing.

Later on, with growing neuropsychological and neurophysiological evidence, further stages and branches were added and the assumption of the strict linearity of the stage sequence was loosened or given up, particularly in the domain of perceptual stages. Stages were seen to overlap in time, and parallel processing was also considered, at least in certain parts of the information stream. For instance, in vision research a distinction between two separate pathways, or information streams, has been proposed based on a number of neuropsychological and neurophysiological findings, and, as a consequence, the notion of a linear change of processing stages has eventually been replaced by the notion of parallel and distributed processing (Desimone & Ungerleider 1989; Milner & Goodale 1995). After an initial common "low-level" feature analysis in the striate cortex (V1) two separate "higher-level" analyses are performed in parallel in the parietal and temporal lobes. The ventral pathway (the "what"-pathway in the temporal lobe) is seen as crucial for cognitive performance like identifying objects (inferotemporal cortex via V4), whereas the dorsal pathway (the "how"- or "where"-pathway in the parietal lobe) is seen as crucial for orientation performance like locating objects (mediotemporal cortex). Milner and Goodale (1995) associated this pathway with the visual guidance of actions (see, e.g., Merigan & Maunsell 1993; Rossetti & Revonsuo 2000, for considerable criticism). Taken together, this results in an elaborate chart with two parallel streams of sequential stages that serve different perceptual functions and can be assigned to different anatomical structures in the brain.

The standard picture of perceptual information processing that has emerged over the last decades can be summarized as follows: When a stimulus is presented, a number of parallel and serial operations is triggered. Some "early" operations are exclusively driven by input information derived from the stimulus, while the "later" ones are controlled by input information from both stimulus and memory information. The end product of perceptual processing – the percept – eventually emerges from combining the distributed outputs of the various operations involved. In any case, there is no action in perception: Perception comes first, and only then comes action (if it comes at all), and there is no direct way they could speak to each other.

**2.1.2. Selection views: From action to perception.** It was certainly not by accident that work on attention had a fresh start with the advent of the information-processing approach (for historical reviews, see Neumann 1996; Van der Heijden 1992). Attentional mechanisms account for the various limitations observed in human performance. Thus, they are assumed to enhance or to inhibit the activity in the various streams and at various stages in the flow of information. The details of attentional theories are assumed to depend on the details of the supposed cognitive architecture.

*2.1.2.1. Capacity limitations.* Correspondingly, early theories of attentional mechanisms completely adopted the linear stage view of information flow. They made an effort to cover various processing limitations, as had become apparent in dual-task experiments. For instance, in the dichotic listening paradigm (Cherry 1953) listeners who are asked to reproduce two dichotically presented messages can usually report only one of them. The filter theory offered by Broadbent (1958; 1971) to capture these and further findings relied on two major assumptions: First, that there is a basic capacity limitation which is due to limitation in the rate of information transmission at the perceptual stage, resulting in a central bottleneck in the flow of information, and, second, that access to the bottleneck is mediated by a filter-like selection mechanism that rejects certain stimulus features from further processing.

Later on, with more progress in experimental research on selective attention tasks, this view became progressively extended: Single-channel conceptions were contrasted with multiple-channel conceptions (e.g., Pashler 1989; Welford 1980), early-selection with late-selection mechanisms (Broadbent 1958; Deutsch & Deutsch 1963), capacity-demanding with capacity-free processes (Shiffrin & Schneider 1977), and specific with unspecific capacity limitations (Kahneman 1973; Navon & Gopher 1979). Up to the early 1980s, one main doctrine remained amazingly constant, however: that capacity limitations are an inherent feature of the processing system and that selection mechanisms are required to handle these limitations. Capacity limitation was the central concept, and selection was considered its functional consequence. This doctrine had important implications for the study and the conceptualization of attentional mechanisms (Neumann 1987; 1996). However, over the last decades this doctrine was challenged by two alternatives at least: the selection-for-action view and the premotor view of attention.

*2.1.2.2. The selection-for-action view.* According to this view, selection came to the fore and lost its role of being just a (secondary) consequence of (primary) capacity limitations. One reason is that empirical research interests have shifted from dual-task paradigms to studies emphasizing selection, like in visual search (e.g., Treisman & Gelade 1980) or spatial cueing (e.g., Posner 1980). In the wake of this shift the functional relationship between capacity limitations and selection has become reversed and eventually replaced by an


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

action-based approach: Selection was no longer seen to follow from system-inherent capacity limitations, but to cause such limitations instead – and to do so for the sake of goal-directed action (cf. Allport 1987; Neumann 1987; 1990; Van der Heijden 1992).

The basic principle behind this view is that any integrated action requires the selection of certain aspects of environmental information (that are action-relevant) and, at the same time, ignore or reject other aspects (that are action-irrelevant). In this sense, studies on attention take their way from action planning to perception: Action planning requires selection that, in turn, modulates perception. The requirements of action control force the system to limit its processing and, thus, make it virtually indistinguishable from an intrinsically capacity-limited system.

The selection-for-action view speaks to the issue of what gets selected and which criteria are involved. It is explicit on what information is picked up and selected, but it has not much to say about the representational structures underlying the mediation between action planning and perception and maintaining action-dependent selection criteria. Whatever the nature of these mediating structures may be, they must allow for fast and efficient interaction between stimulus-related and action-related processing.

*2.1.2.3. The premotor view of attention.* This view also suggests that selection neither results from nor requires an attentional control system separate from action-perception cycles. Rather, selective attention derives from an exogenous or endogenous activation of spatial cortical maps, in which spatial information is transformed into movements. As a consequence of activation, there is an increase of motor readiness and a facilitation of stimulus processing at locations toward which the motor program is directed (Rizzolatti & Craighero 1998; Rizzolatti et al. 1994). The "premotor" connotation of the theory results from the assumption that it needs only the preparation of the motor program (the process we call "action planning") but not necessarily its execution. This allows integration of observations of covert attentional orienting without the overt action component, for example, when attention is shifted to a stimulus in the periphery during eye fixation.

Originally, the premotor view of attention was only applied to those cortical maps that code space for oculomotor behavior (Rizzolatti et al. 1987). Similar ideas of a close link between oculomotor behavior and covert attention shifts have been discussed before, for example, in assuming that attention shifts precede each goal-directed saccade and its programming (Klein 1980; Posner & Cohen 1980; 1984; Rayner 1984). Nowadays, the premotor view is extended to any map that codes space, that is, to maps that control movements of the head, the arm, or other body parts (Rizzolatti & Craighero 1998; Rizzolatti et al. 1994). Nevertheless, the view addresses exclusively and is restricted to phenomena of *spatial* attention. It does not speak to how the requirements of action planning interact with the processing of other stimulus dimensions than space and it has not much to say about the representational structures underlying the mediation between action planning and perception.

**2.1.3. Adaption views: Perception and action.** Another interesting perspective on the relationship between perception and action can be found in the vast literature on spatial and temporal orientation and adaption (cf., e.g., Andersen 1988; Colby 1998; Howard 1982; Howard & Templeton 1966; Paillard 1991; Redding & Wallace 1997). This perspective refers to the notion that perception and action control make use of shared reference frames with respect to space and time. In particular, it posits shared frames for environmental objects and events and for the actor's body and his or her movements. These reference frames serve to specify the spatial and temporal relationships between environmental events and bodily actions and, thus, to coordinate the one with the other. Though this assumption appears to be a prerequisite for the successful functioning of sensorimotor systems it is not often explicitly stated in the literature – perhaps because it is too evident (see, however, Brewer 1993; Grossberg & Kuperstein 1989; Howard 1982; Redding & Wallace 1997). If no such shared reference frame existed, one could not explain how animals can reach for a target at a perceived location or catch a flying ball at the time and place of expected collision (Lee 1976).

The principle of shared reference frames has two faces: unity and diversity. The notion of a fundamental unity between perceptual space and action space is derived from functional considerations. Animals are spatially extended systems that move around in their physical environment. An efficient way to successfully interact with the environment is to furnish them with a representational system that allows representation of the Where and the When of both environmental events and body-related events (actions) in a shared spatio-temporal reference frame. Such a shared representation is reliable in the sense that it allows us to coordinate represented actions and represented events (e.g., intended actions with perceived events). Moreover, in order to be valid, this representational system needs to contain veridical information in the sense that it recovers the spatial and temporal pattern of happenings in the distal environment (cf. Brewer 1993; Prinz 1992).

Recent literature, however, emphasizes diversity over unity at both the functional and the structural level (Jeannerod 1983; Paillard 1991; Redding & Wallace 1997). For instance, in the functional analysis of visually guided movements one has to distinguish, on the sensory side, between a number of independent spatial maps, or reference frames (e.g., eye-, head-, trunk-related) and transformations mediating between them. The same applies to the motor side where movements can be specified in a number of reference frames, too (e.g., joint-, effector-, trunk-related ones). Electrophysiological recordings have also stressed diversity from a structural point of view, suggesting multiple representations of spatial information in a number of maps at different cortical sites (e.g., Andersen 1988; Colby 1998; Ohlsson & Gettner 1995). Some of these representations can be specified with respect to the body parts to which they are anchored (like eye, hand, arm) or with respect to the action patterns to which they contribute (like grasping, reaching). Some others can even be specified in terms of body-independent, allocentric coordinates (such as object-relative movement directions).

Given such modularity of spatial information processing in both functional and structural terms, it is not too surprising that performance dissociations can be observed in a number of experimental tasks (e.g., Aschersleben & Müsseler 1999; Bridgeman et al. 1981; Fehrer & Raab 1962), particularly with respect to the distinction between the dorsal and the ventral stream of processing (cf. Desimone & Ungerleider 1989; Milner & Goodale 1995). We must not



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

forget, however, that these rare examples of diversity and dissociation are exceptions from the normal case of unity and association and, thus, integrated action (Milner & Goodale 1995, ch. 7; Rossetti & Revonsuo 2000). Unless we understand how the brain solves the problem of the binding of distributed representational activities (across maps and across units within maps) we have no way of understanding how unity and diversity can be combined and how one emerges from the other (Singer 1994; Treisman 1996).

Yet, despite the intricacies of the functional architecture and the complexity of the transformational computations involved, there can be no doubt that the system eventually uses corresponding frames of reference for perception and action planning. This functional unity works not only from perception to action – as evident from the high degree of precision with which animals fix the spatial and temporal details of their actions on the basis of the information provided by their perceptual systems – but from action to perception as well. As animals move, perceptual information cannot be interpreted in an unambiguous way without reference to action-related information. This, in turn, requires that these two pieces of information interact with each other within the same reference system. Such cross talk can occur at two levels, compensation and adaptation.

The notion of compensation refers to the fact that in order to interpret any change in the spatial distribution of signals at their receptor surfaces, animals must have a way to compensate for their own body movements. In other words: the system has to take the animal's body movements into account before it can use the sensory signal to recover the structure of the environmental layout (cf., e.g., Bridgeman 1983; Epstein 1973; Matin 1982; Shebilske 1977). For instance, according to the classical reafference model, this is performed by a subtractive operation by which perception (the retinal location signal) gets canceled for the effects of action (saccadic eye movements). Clearly, a subtractive model like this implies at least commensurate (or even identical) representations on the perception and action side (von Holst & Mittelstaedt 1950; MacKay & Mittelstaedt 1974).

The notion of adaptation refers to the flexibility of sensorimotor couplings and to the fact that perception can, within certain limits, be educated by action planning. For instance, in studies of distorted vision (Kohler 1964; Redding & Wallace 1997; Stratton 1897; Welch 1978), it has been shown that such education can work either way: Perception may teach action and action may teach perception at the same time, again suggesting commensurate or identical representations on both sides (see also Koenderink 1990; Van der Heijden et al. 1999a; Wolff 1987; 1999).

In sum, we conclude that the notion of commensurate or even identical representations and shared reference frames for perception and action is widespread in the literature on spatial orientation. It appears to be a natural notion that requires no explicit defense, and this remains to be the case even in the light of what is known about multiple representations of space in the brain.

**2.1.4. Challenges for a cognitive theory of perception.** Although the functional autonomy of perceptual processes is implicitly or explicitly presupposed in most theoretical approaches to perception, it has been questioned from time to time. In fact, a number of studies have demonstrated that there is more interaction between perception and action planning than the standard framework would allow.

*2.1.4.1. A role for action in perceptual processing.* In the literature different accounts can be found that seem to indicate some impact of action-related processes on perception. Already in the late nineteenth century Lotze introduced a prominent motor theory of perception with the "theory of local signs" (Lotze 1852; for an historical overview see Scheerer 1984). It basically claimed that space perception arises from a combination of two sources of information. He assumed, first, a qualitative map of visual sensation and, second, a quantitative map of metric necessary to fovealize an object. What is perceived was assumed to result from the qualitative visual sensation "enriched" by the quantitative map of oculomotor behavior (or vice versa; for modern versions of this idea, see Koenderink 1990; Van der Heijden et al. 1999a; Wolff 1999). These notions are used, for example, to account for phenomena of visual localization (e.g., Müsseler et al. 1999; Van der Heijden et al. 1999b).

While in the early motor theories motor processes fulfill a representational function (i.e., they provide the system with a given metric – an assumption difficult to maintain, cf. Scheerer 1984), modern views assign to them an adaptive function. What is perceived is influenced by previous motor adjustments and, at the same time, it is a precondition for future adjustments (see also below and, for an overview, Welch 1986). However, nearly all motor theories deal exclusively with phenomena of space perception. The most prominent exception is the motor theory of speech perception (e.g., Liberman 1982). We will deliberately exclude this example, partly because the language domain exceeds our scope here, partly because empirical support seems weaker than the elegant theoretical principle suggests (cf. Jusczyk 1986; Levelt 1989, ch. 111).

Another field where interactions between perception and action can be studied is the perception of human action (Cutting & Proffitt 1981; Freyd 1987; Johansson 1950; 1973; for an overview see Stränger & Hommel 1996). Much of the evidence from this work can be summarized by concluding that the way people perceive other people's actions appears to rely on specific representational structures that contain information that goes far beyond the information provided by the actual stimulus. For instance, when people are confronted with traces of handwritten letters, drawings, or body movements they are often capable of inferring the kinematics or even the dynamics of the movements by which these traces were generated (Babcock & Freyd 1988; Freyd 1987; Kandel et al. 1995; Runeson & Frykholm 1981). These studies suggest that the visual perception of actions and their consequences may draw on action-related representational structures subserving both the generation and perception of action patterns.

Related evidence comes from studies showing that semantic judgments about actions are facilitated if preceded by performing hand gestures that match the to-be-performed action in certain respects (Klatzky et al. 1989). A further example is provided by studies on apparent biological motion (Heptulla-Chatterjee et al. 1990). These studies demonstrate that when apparent motion is induced by a stimulus pattern in which the human body forms a constituent part the resulting motion does not always follow the principle of the shortest path (as it would usually do with physical objects). Instead, as if to avoid anatomically impossible movements, the motion takes longer paths and detours.

No less dramatic is the implicit impact of action on perception in a series of elegant studies by Viviani and his col-


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

leagues. As had been shown in earlier studies, the velocity of drawing movements depends on the radius of the trajectory's curvature (Viviani & Terzuolo 1982). Interestingly, the same lawful relationships also seem to be effective in perception (Viviani & Stucchi 1989; 1992; Viviani et al. 1997). For instance, the velocity of a moving dot seems to be uniform if (and only if) it actually follows the law governing movement production (similar effects can also be found with linear motion; see, e.g., Mashhour 1964; Rachlin 1966; Runeson 1974). This suggests that production-related knowledge is implicitly involved in perceptual processing – at least as far as the perception of actions and action effects is concerned.

Very recently, support for shared mechanisms for action perception and action control has also been provided by neurophysiological studies (for an overview, see Decety & Grèzes 1999). For instance, Rizzolatti and his group describe "mirror neurons" in the premotor cortex of the monkey. These neurons are active both when the monkey performs a given action and when it observes a similar action performed by the experimenter, which typically must involve an interaction between the agent and an object (Gallese et al. 1996; Di Pellegrino et al. 1992). Similar findings were obtained in PET-studies with human participants when execution, observation, and imagination of grasping movements were compared (Rizzolati et al. 1996; Grafton et al. 1996), suggesting a shared representational basis for perception and action planning. The same conclusion is suggested by a study in which transcranial magnetic stimulation was applied during action execution and action observation (Fadiga et al. 1995).

*2.1.4.2. A role for action in perceptual learning.* For obvious reasons the study of perceptual learning has seen more discussion about interaction between action and perception. Learning is usually believed to result from actions and their outcomes, and therefore the study of perceptual learning requires examining long-term consequences of action on perceptual processing (for overviews see Gibson 1969; Hall 1991). While the major part of the literature covers issues of spatial orientation, some part is also devoted to the emergence and improvement of perceptual discrimination in the domain of patterns, objects, and events. However, there appears to be more theory than solid evidence.

One of the classical theoretical positions has become known as the differentiation theory of perceptual learning (Gibson 1969; Gibson & Gibson 1955). According to this theory, perceptual learning is the art of picking up, or extracting, those invariants of stimulation that are suited to directly specify certain body movements or actions. From its beginning this theoretical program was anticognitivist (or, historically speaking, anti-Helmholtzian) because it relies on the extraction of action-relevant information (presumably contained in the distribution of stimulus energy) rather than computational operations generating internal codes and action programs. The theory had initially been developed to account for the psychophysics of spatial orientation (Gibson 1950) and was later broadened to account for the ecology of perception in general (Gibson 1979; cf. Turvey 1977; Fowler & Turvey 1982). Central to the theory is the notion of affordances which, roughly speaking, stand for specific combinations of objects and events as taken with reference to their functions for the animal – typically with reference to actions of certain kinds.

Unlike differentiation theory, which believes that actions act back on the early perceptual mechanisms of feature extraction, the theory of associative enrichment posits that perceptual learning occurs at a much later stage. The basic assumption is that when a given stimulus is frequently coupled with a given response, the information derived from that stimulus will become associatively enriched with response-produced cues that then will help to discriminate this stimulus from other ones coupled with other responses (acquired distinctiveness of cues; cf. Miller & Dollard 1941; Postman 1955). Such response-produced cues could come, for instance, from immediate proprioceptive byproducts of the response as well as its more remote effects.

The ecological approach has been quite successful in accommodating for a number of findings in perceptual learning and development (cf. Gibson 1969; Thelen & Smith 1994, Ch. 8/9) – more, perhaps, than the enrichment approach which has seen only partial support from experimental evidence (cf. Arnoult 1957; Cantor 1965; Hall 1991). Still, the issue appears to be unsettled. One reason may be that the differentiation theory is broader and less specific than the enrichment theory and can therefore easily account for a variety of facts (and, if necessary, be adapted to accommodate new findings). In fact, it is not easy to imagine how it could be falsified at all. Moreover, the truth might lie in a combination of the two theories. Obviously, instead of excluding each other on any logical grounds, they share the common belief that action plays an important role in teaching perception in a specific way. So far, however, they both have not spelled out the details of this teaching procedure in a satisfactory way.

### *2.2. Views on action planning*

In the past, theorizing about action has been organized in two separate lines of thought and been shaped by two different conceptual frameworks, namely the *sensorimotor view* and the *ideomotor view.* Theories from the sensorimotor approach tend to view actions as responses to stimuli, that is, as following from external causes. Conversely, theories from the ideomotor approach tend to view actions as following from internal causes like goals or goal-related cognitions. In this section we argue that a comprehensive account of the cognitive foundations of action requires combining the two views in a new way.

**2.2.1. Sensorimotor views: From stimuli to responses.** In the sensorimotor view of action, actions are regarded as "reactions," that is, as responses triggered by stimuli. Strict versions of the approach (like classic behaviorism) claim that such reduction to stimulus conditions is a necessary and at the same time sufficient condition for a full account of action. More liberal versions may also consider a role for additional factors that cannot be traced back to the actual stimulus situation.

Historically, the sensorimotor line of thought has been the mainstream in action theory for decades, if not centuries. This is true of both physiology and psychology. One of the early influential sources to which its origin can be traced back is Descartes' (1664) analysis of the relationship between perception and action. According to Descartes, actions must be understood to be the result of the perception of events. This doctrine has laid the groundwork for the notion of the sensorimotor arc, which ever since has deeply



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

influenced psychological theorizing on action. This can be shown in a number of independent historical and theoretical contexts such as, for example, the foundation of reaction time measurement (e.g., Donders 1868/1969), the formation of the behaviorist program or, more recently, the development of the linear stage theory of human performance (e.g., Massaro 1990; Sanders 1980; 1998; Theios 1975). The fact that sensorimotor theories of action have so far been clearly more successful than ideomotor theories may be rooted in the difference between the explanatory strategies of the two views and their methodological implications. Unlike the ideomotor view, which explains actions in terms of mental causes, the sensorimotor view refers to physical causes. As a consequence, the sensorimotor framework offers a tool for both the study and the explanation of action in physical terms. This dual advantage may explain part of its historical success.

Theories of action address two major systematical problems, learning and performance. Learning theories from the sensorimotor domain tend to adopt a structural stance, explaining how associations between stimuli and responses (or their internal representations) are created and how their strength depends on factors like contiguity, reinforcement, and practice. Further, they address the issue of how these associations interact in networks that link large sets of responses to large sets of stimuli. Conversely, theories of performance tend to adopt a functional stance, modeling the chain of processing stages and operations by which stimulus information is translated into motor commands, to the effect that responses are selected on the basis of the information available.

These theories refer to experimental task contexts with well-defined sets of stimuli and responses (such as instrumental conditioning, paired-associate learning, or disjunctive reaction tasks). In these tasks, stimuli and responses are individuated on the basis of the task, including the instructions. Task and instructions specify (1) which stimuli can occur, (2) which responses can be selected, and (3) which rules govern the mapping of responses to stimuli. The core assumption shared by all brands of sensorimotor theories of action may be called the stimulus trigger hypothesis. It holds that, at least in such well-structured task contexts, the presentation of the stimulus is both a necessary and sufficient condition for triggering the appropriate response. Accordingly, any action is a re-action: it comes into being as a causal consequence of stimulation.

**2.2.2. Ideomotor views: From goals to movements.** In contrast to sensorimotor views of action, ideomotor views stress the role of internal (volitional) causes of action and at the same time disregard the role of external (sensory) causes. Correspondingly, the evidence ideomotor theories are grounded on does not come from stimulus-controlled reaction tasks but rather from more open situations where individuals pursue certain goals and, from time to time, perform certain actions in an attempt to approach or achieve them. In this view, actions are considered creations of the will – events that come into being because people pursue goals and entertain intentions to realize them.

Historically, ideomotor views to action have entered the scientific discussion much later than sensorimotor views. This is obviously due to the fact that they deal with latent, internal rather than manifest, external causes of action. Moreover, the role of these causes appears to be different. Unlike a stimulus that may be regarded as a direct cause for the responses to follow, a goal appears to be a causal determinant that, at first glance, seems to work backward in time. Lotze (1852), Münsterberg (1888), and James (1890) were the first to solve the puzzle, at least in principle. Their solution was based on a distinction between the goal state itself (as achieved through the action – and, thus, following it in time) and its cognitive representation (as formed before the action – and, thus, potentially involved in its causation).

Though this move solved the puzzle of backward causation it did not help much in bringing goals to the forefront of action theory and assigning to goals (or their representations) the same status and dignity that everybody (and every theory from the sensorimotor domain) would find natural to assign to stimuli (or their representations). Again, what remains is a difference in methodological status: Stimuli are directly observable entities, and therefore stimulus information is easy to record and/or manipulate. Goal representations, however, are unobservable entities that cannot be recorded and manipulated that way. If there is no solution to this problem, the study of the role of goals must, for principled reasons, be much more delicate than the study of the role of stimuli for action. As a consequence it is not too surprising that we do not possess a comprehensive conceptual framework for understanding the role of goals in action.

This is not to say that goals have been completely ignored in action research. First, in the motor control literature, a substantial body of knowledge has been accumulated on target-related performance such as goal-directed hand movements (cf. Jeannerod 1988; 1990) and saccadic eye movements (cf. Carpenter 1988; Kowler 1990). Here, targets may be viewed as goals that are specified in terms of spatial and temporal motor coordinates. Yet, in most of these paradigms the targets (or some aspects thereof) can be perceived in the same way as the triggering stimuli. Therefore this literature is only of limited use for ideomotor theories of action whose main point is to clarify the role of internal goals.

Second, the study of goal selection and goal implementation has ever been one of the key issues in theories of motivation and volition (e.g., Ach 1905; Bargh 1996; Gollwitzer 1996; Lewin 1926; Mischel 1996). These theories have much to say about the internal dynamics of processes that precede the action. Yet, most of them are virtually silent when it comes to the details of the operations by which goal representations contribute to action planning proper.

Third, more general theories are available on the nature of the mechanisms for goal-directed control of cognition and action. For instance, as far as cognition is concerned, production-rule-based computational models like ACT-R (Anderson 1993) or EPIC (Meyer & Kieras 1997) provide a central role for goals and goal representations in their architecture. Conversely, as far as action is concerned, there has been a long-standing debate on the role of goal representations in learning and performance, that is, on how people learn to build up representations of future goals from consequences of previous actions and how these goal representations later become involved in action control (e.g., Ach 1905; Greenwald 1970; Hull 1952; Konorski 1967; Miller et al. 1960). Still, none of these approaches has provided a sufficiently detailed framework for the mechanisms underlying the formation and operation of goal representations in action control – sufficient in the sense that they can


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

be spelled out in terms of the involved cognitive structures and mechanisms.

In summary, it seems that goals have not lost much of the elusive character they used to have in the times of Lotze, Münsterberg, and James – at least as far as their role for action control is concerned. Ideomotor theories of action are much less developed and elaborated than sensorimotor theories with respect to both performance and learning.

In the ideomotor domain, theories of learning tend to adopt a structural stance, too, explaining the formation of linkages between body movements and their consequences. In order to lay the ground for the explanation of voluntary action these linkages must meet two requirements (cf. Prinz 1997a). First, following James' (1890) analysis of ideomotor action, these linkages need to include both resident and remote effects. Resident effects refer to the mandatory reafferent consequences that go along with performing certain body movements. The perception of resident effects is therefore based on body movements, for example, an arm movement required to throw a ball. Remote effects refer to the various environmental consequences that may follow from performing certain movements. The perception of these effects is based on environmental events, like, for example, a ball bouncing against a wall. Second, though the learning refers to linkages between movements and their effects, the result of this learning needs to be organized in a way that allows use of the linkages the other way round, that is, to go from intended effects to movements suited to realize them.[1]

This is where performance theory of voluntary action comes into play. If one takes for granted that the links between movements and effects can be used either way, a simple conceptual framework for the functional logic of voluntary action offers itself. This framework suggests that actions may be triggered and controlled by goal representations – that is, representations of events the system "knows" (on the basis of previous learning) to be produced by particular movements.

The core assumption shared by various brands of ideomotor theory may be called the "goal trigger hypothesis." It holds that goal representations that are functional anticipations of action effects play a crucial role in action control (Greenwald 1970; 1972; James 1890; Lotze 1852; Prinz 1987; 2002). For instance, Lotze speaks of "Vorstellungen des Gewollten" (image of the intended events; Lotze 1852, p. 301) and James of the "bare idea of a movement's sensible effects" which serves the function of a "sufficient mental cue" to the movement itself (James 1890, p. 522). Given this crucial role for goal representations it is natural that, according to ideomotor theory, the proper way to individuate actions is neither in terms of stimuli nor responses but in terms of goals and goal representations.[2]

### 2.2.3. Challenges for a cognitive theory of action planning

*2.2.3.1. Combining the two views.* Obviously, a full account of the cognitive foundations of action control requires combining the two lines of thought. Action depends on both external and internal causes, and this needs to be reflected in a combined theoretical framework.

*2.2.3.1.1. (Most) Reactions are voluntary actions, too.* Consider, as an example, a disjunctive reaction task where, on each trial, the participant selects one out of several reactions in response to the presentation of one of several stimuli. As was indicated above, performance in such tasks is usually analyzed as if response selection is more or less fully determined by the stimuli presented. This, however, is obviously not the case: The presentation of the stimulus is necessary, but it is certainly not a sufficient condition for the response to occur. Nothing will happen upon stimulus presentation until the participant has been instructed to respond in a particular way and he or she is willing to do so. In other words, in order for a response to occur, two conditions must be fulfilled: There must be an appropriate stimulus and an appropriate intention to respond to that stimulus in a particular way (Hommel 2000a; 2000b). This may be trivial to state – and everybody knows it – but theories in the sensorimotor domain have never really recognized this fact. It is not the stimulus that presses the response key in the reaction task. What is rather required as underlying causal structure is a disposition to press a particular key under particular stimulus conditions.

*2.2.3.1.2. Voluntary actions are reactions, too.* Consider, on the other hand, voluntary actions like opening a door or going to the movies. For these examples, too, it is trivial to state that the action is not only dependent on internal causes (goals) but on external causes as well. Whatever the goals in these examples may be, they do not by themselves specify the detail of the actions suited to realize them. All of these details need to be specified by taking the actual pattern of a number of external factors (stimuli) into account. Motivational theory has to some extent given recognition to this fact – for instance, in the claim that, in order to perform voluntary actions, internal intentions need to interact with external opportunities – to the effect that both are required to realize the action (Lewin 1926). Again, it is obviously not the goal itself that opens the door but rather a disposition to realize a certain goal under certain circumstances and to fix the details of the action in response to the details of the circumstances.

*2.2.3.1.3. Meeting the challenge.* In summary, then, in order to account for stimulus-triggered response performance, we need to develop a novel conceptual framework that treats reactions as goal-directed actions. Actions and reactions need to be regarded as segments of body movements that are individuated on the basis of goals – for instance, like in throwing a ball in order to hit a target or in pressing a key in order to switch the light on. In other words, actions are, by definition, structures that link movements to goals – and vice versa (Prinz 1997b). Of course, it may not always be easy to conceptually distinguish movements and goals. The distinction is easy to apply to cases where the goal lies beyond the movement and is clearly distinct from it, like in the two examples just mentioned. It is less obvious how it can be applied to cases where the goal appears to reside in the movement itself, like in pressing a response key in a laboratory task. Still, there remains a difference: there is a movement first (finger and key going downward) and a goal state resulting from that movement second (the key being pressed down).

This view opens an interesting perspective not only on goals themselves, but also on the interaction between goals and stimuli as internal versus external causes of actions. With this view, goals and stimuli are commensurate, as they both refer to environmental events – events that are going



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

on and are being perceived in the case of stimuli and events that are being planned and to be effected in the case of goals. As we will discuss in the section "The Theory of Event Coding," one of the central assumptions of TEC will indeed refer to the notion of a common representational domain for stimuli and goals, or perceived and intended events, respectively.

*2.2.3.2. Action induction.* One of the major empirical challenges that a theory of the cognitive foundations of action control has to meet comes from action induction. By this term we refer to a variety of observations suggesting much closer functional links between perception and action planning than the standard frameworks provide – links that appear to be based on inherent similarity between perception and action planning rather than on acquired arbitrary connections.

*2.2.3.2.1. Imitative action.* There is a large variety of imitative actions, ranging from seemingly simple imitations of elementary orofacial gestures in newborns and infants (e.g., Meltzoff & Moore 1977) to instances of observational learning of habits, attitudes or even traits in adults (Bandura 1977; Bandura & Walters 1963). Correspondingly, a number of functional distinctions have been proposed in the literature (Scheerer 1984). Still, there is one feature common to all varieties of imitative action; namely, that the imitator's action resembles the model's action in one or the other respect. In the context of social learning this resemblance may refer to the action's eventual outcome (like wearing the same clothes as the model), whereas in the context of skill acquisition it may refer to the kinematics of the movement pattern itself. Theories of imitation have treated the issue of resemblance in two different ways (Prinz 1987). One way is to consider it a by-product emerging from the operation of mechanisms that are based on contiguity and reinforcement rather than similarity (e.g., Gewirtz & Stingle 1968). The other way is to consider similarity the functional basis of imitation – functional in the sense that imitative action occurs *by virtue of* similarity between perception and action (e.g., Meltzoff & Moore 1977; Piaget 1946). In any case, the theoretical challenge inherent in the ubiquitous occurrence of imitative action is still unresolved. How is it possible that the mere perception of certain actions in another person can give rise to similar actions in the perceiver? How can similarity be effective between what s/he perceives and what s/he performs?

*2.2.3.2.2. Sympathetic action.* When a person watches a scene in which he or she is deeply involved, sympathetic actions may sometimes be observed in this person – sympathetic in the sense that they are clearly related to the happening in the scene and appear to be induced by them. Sometimes the term ideomotor action has also been used to denote a particular class of induced actions. Yet, as this term has also been used in a much broader sense (Carpenter 1874; James 1890), we suggest the term of sympathetic action in order to avoid confusion (Prinz 1987). Sympathetic action is clearly different from imitative action. First, to quote a distinction from Katz (1960), sympathetic action tends to be synkinetic (an immediate on-line accompaniment of perception) rather than echokinetic (a delayed offline follower). Second, sympathetic movements often occur without, sometimes even against the spectator's will. Third, despite their involuntary character, they appear to be strongly dependent on his/her intentional involvement in the scene being watched, suggesting the seeming paradox of an involuntary action that is still under intentional control. Until recently, the details of the relationship between the happenings in the scene and the sympathetic movements accompanying them have not been systematically studied (see Knuf et al. 2001). Whatever this relationship may be – sympathetic action also suggests closer links between perception, action, and (tacit) intention than the classical frameworks assume.

*2.2.3.2.3. Synchronous action.* When exposed to rhythmic sounds many listeners find it easy to dance or carry out other rhythmic movements in accordance with the sound pattern, and some will even find it hard to suppress such synkinetic accompaniment of the sound pattern they are exposed to. As Fraisse (1980) has pointed out, the action-inducing power of such auditory-motor synchronizations is particularly strong, suggesting a special and privileged link between perception and action. In movement synchronization the action-inducing power of perception is restricted to the temporal domain. The perceptual event itself does not specify which movements to perform and which limbs to use. However, once these choices have been made by the listener the timing of the movements is captured by the structure of the sound pattern – provided that it exhibits sufficient regularity to allow for anticipation of temporal structure. Synchronized action is related to both imitative and sympathetic action. With imitative action it shares the feature that the action's structure is modeled after the structure of the perceptual event. With sympathetic action it shares the features of synkinetic accompaniment and spontaneous, involuntary occurrence. Again, the resemblance between perception and action suggests that one induces the other by virtue of similarity.

*2.2.3.2.4. Compatible action.* Spatially compatible action can be considered the spatial counterpart of synchronized action. Effects of spatial compatibility between stimuli and responses can be shown in simple experimental demonstrations. A simple example is provided by a choice reaction task with two stimuli and two responses where on each trial a stimulus light is flashed to the left or to the right of a fixation mark and one of two response keys is operated in response to the stimulus, either the left one or the right one. A setup like this allows for two tasks differing in the assignment of responses to stimuli. When the assignment is spatially compatible, stimuli and responses share a common spatial feature (both left or both right) – in contrast to the incompatible assignment where they will always exhibit two different spatial features (right-left or left-right). As has been shown in a large number of experimental studies, response performance for compatible assignments is clearly superior to incompatible assignments (in terms of response times and error rates; cf. Fitts & Seeger 1953). Though it may not be impossible to account for the effect in terms of practice and contiguity, it is more natural to suggest that responses can be pre-specified by stimuli on the basis of shared features (e.g., Greenwald 1970; Kornblum et al. 1990). This is not far from concluding that perception induces action by virtue of similarity – and from raising the question of how such induction may occur.


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

### 2.3. Views on perception and action planning

Obviously, then, in order to achieve a deeper understanding of the mechanisms of perception and action planning, we need to come up with a more integrated approach, recognizing the intimate relationships between perception, cognition, and action planning. In other words: We need to argue against what MacKay et al. (1987) have called the separate-and-unequal approach to perception and action – a view that has been challenged several times (e.g., Decety & Grèzes 1999; Gallese et al. 1996; MacKay 1987; MacKay et al. 1987; Neisser 1985; Turvey 1977; Viviani & Stucchi 1992; von Hofsten 1985) but still dominates the theoretical discussion.

On the one hand, theories of perception need to meet the challenge from various forms of interaction between perception and action planning in processing and learning. On the other hand, theories of action planning need to meet the challenge of providing roles for similarity between perception and action planning as well as the operation of goals in action control. We believe that the time has come to meet these challenges. Recent evidence from a number of fields has provided support for strong linkages between perception and action, for example, brain imaging (Decety, in press; Jeannerod 1997: 1999; Passingham 1993), single cell recording (Di Pellegrino et al. 1992; Rizzolati et al. 1996; Perrett et al. 1989), executive functions (Allport et al. 1994; Monsell & Driver 2000), voluntary action (Bargh & Barndollar 1996; Gollwitzer 1996; Hershberger 1989), imitation (Meltzoff 1995; Meltzoff & Prinz, in press; Nadel & Butterworth 1999; Prinz, in press), and conditioning (Rescorla 1991). Before we unfold our own ideas on these matters (sect. 3) we will briefly sketch two major views on related issues.

In a way, interactions between perception and action are at the heart of the ecological approach to perception and action, as advanced by Gibson (1977; 1979; cf. Michaels & Carello 1981; Reed 1982; 1996). According to this approach, a particular role in the mediation between perception and action is played by the notion of affordances. Affordances specify aspects of the environment with reference to the animal's body and its action capabilities, like, for instance, the "climbability" of a staircase and the "sittability" of a chair. A further claim is that the invariants that inform the animal about the action potential of its environment are inherent in the stimulus array. As a consequence, there is no need for any elaborate processing of the stimulus information. Instead, the animal's option is to take or leave the invitations inherent in the actual affordances. When it decides to take one of them, action gets tuned to perception, and the details of the upcoming action get directly specified by the information picked up from the environment (e.g., Reed 1993).

Ecological approaches have introduced important novel perspectives to the study of perception and action. One is that they take the animal's body and its actions into account as a reference for both the detection of affordances on the perceptual side and their use for the formation of coordinated movement patterns on the action side. At the same time they stress the importance of action for perception, that is, the instrumental role of movements for the pick-up of complex invariants of stimulation. Another important novel perspective is that they adopt a realistic and, thus, anti-computational stance, believing in direct (Gibsonian) detection of information rather than indirect (Helmholtzian) processing.

Cognitive views on perception and action differ from ecological views with respect to both scope and explanatory principles. As regards scope, cognitive approaches focus on the contributions of memory systems and stored knowledge to perception and action. Unlike ecological approaches, which emphasize the role of affordances through which action gets directly tuned to perception, cognitive approaches consider the role of instructions and intentions for the formation and implementation of task-specific cognitive dispositions, or task sets. What they try to explain is action planning, that is, the individual's ability to select, prepare, and initiate arbitrary voluntary actions in response to arbitrary environmental events, and to do so on the basis of rules that may even change from moment to moment.

As regards explanatory principles, cognitive approaches differ from ecological approaches in two basic ways. First, instead of individuals' physical bodies and their action capabilities, they consider their knowledge structures as the central reference for both perceptual analysis and action planning. Second, and related to this, they adopt a constructivist and, thus, representational stance, emphasizing information processing rather than detection.

We hold that a full theory of perception and action will eventually have to speak to the relationships both between affordances and movements and between perceptual events and goal-directed actions. TEC is not meant to offer the full theory. Instead, we focus on core issues inherent in cognitive views, that is, the representation of events and the planning of voluntary actions. For these issues we offer a framework that views perception and action as closely coupled as the ecological approach has claimed for movements and affordances.

## 3. The theory of event coding

We argue that a new conceptual framework is needed for a better understanding of the relationship between perception and action planning, and we believe that TEC offers this framework. TEC is based on the central notion that perception, attention, intention, and action share, or operate on, a common representational domain—a notion we will specify and discuss in this section. In constructing TEC, we have drawn on many ideas from other theoreticians, especially, of course, those emphasizing the intimate relationship between perception and action planning. For instance, we share the general perspective of Dewey (1896) and Gibson (1979) that perception and action are functionally linked and that it is only their coordination that allows for adaptive behavior. We further adopt the notion put forward by Greenwald (1970), James (1890), and Lotze (1852) that action control is anticipatory, that is, controlled by representations of intended action effects. And we also follow Allport (1987) and Singer (1994) in assuming that representations of perceptual contents and action plans are content-specific composites of codes presumably stored in a distributed fashion.

In our view, TEC forms the core of what one may call a (Meta-) Theory of Perception and Action Planning, thus, a framework that would allow a fully integrated view on a substantial part of human information processing. In this section, we will describe the theory as developed so far. In



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

particular, we will give a list of the central assumptions underlying and making up the theory, and describe the basic anatomy of our major theoretical construct: the event code.

### 3.1. Event coding: Basic principles and assumptions

It is fair to say that how stimulus and response representations are related has not been a major concern of information-processing approaches to human cognition. With very few exceptions, the basic idea underlying most approaches is that stimulus codes are formed in some "perceptual" domain and response codes in some "response" or "motor" domain, and without "voluntary," "intentional," or "controlled" translation processes there is not much contact between them (see, e.g., Rosenbloom & Newell 1987; Sanders 1990; Teichner & Krebs 1974; Theios 1975). Only recently, a number of so-called dual-route models have been suggested that allow for both "voluntary" translation of stimulus into response codes and some kind of direct, automatic, and stimulus-induced activation of response codes (e.g., De Jong et al. 1994; Kornblum et al. 1990; for an overview, see Hommel 2000a). However, even in these models the relationship between stimulus and response representations is not described in any detail. In fact, the assumptions are usually restricted to postulate that some kind of interaction *exists* (typically asymmetric effects of stimulus on response processes, not vice versa), but it is not explained why this is so and how it works. As we will show, TEC provides a promising starting point for developing such explanations. But let us begin with discussing the basic principles and assumptions underlying our framework.

**3.1.1. Common coding of perceptual content and action goals.** In contrast to previous approaches to human information processing, we do not share the seemingly obvious (though usually implicit) assumption that perceiving a stimulus object and planning a voluntary action are distinct processes operating on completely different codes. We claim that perceiving and action planning are functionally equivalent, inasmuch as they are merely alternative ways of doing the same thing: internally representing external events (or, more precisely, interactions between these events and the perceiver/actor).

There are obvious objections to our claim: Isn't perceiving a rather passive mode of merely registering the properties and things of our environment that our actions seek to actively change? In our view, this kind of characterization overlooks first, that, perceiving the world is a process of actively acquiring information about the perceiver-environment relationship, including all sorts of movements of eye, hands, feet, and body, particular allocations of attention and other cognitive resources, and so forth, and that, second, action would actually run blind without being perceptually informed about its bodily and environmental preconditions, its progress, and its consequences (Dewey 1896; Gibson 1979). That is, the process of perceiving both presupposes and affords active behavior and performing an action both relies on and produces perceptual information. In that sense, perceptual or stimulus codes and action or response codes all represent both the result of, and the stimulus for, a particular sensorimotor coordination. If so, there is no theoretical reason to draw a conceptual distinction between anticipating a perceptual event and planning an action or between actually perceiving and carrying out an action plan.

Moreover, as already touched on above, voluntary actions can be seen to come into being by anticipating their distal effects (cf. James 1890; Lotze 1852). This also implies that perceived events and action-generated distal events are coded and stored together in one common representational domain (Prinz 1990). Above all, action-generated effects include body-related afferent information, that is, anticipated proprioceptive feedback, but they can also contain visual information about the anticipated position of the arm during and/or after a movement. In motor control literature this view of representing actions in terms of their action goals is widely spread (Jeannerod 1999; Rizzolatti et al. 1997; Viviani & Stucchi 1992). However, TEC is much more radical in that it is based on the assumption that action effects could refer to any kind of response- or action-contingent events (see also Hoffmann 1993; Hommel 1997; Meltzoff et al. 1991). In other words, switching on a light will not only produce a body-related tactile and kinesthetic feedback at the hand, but also an afferent visual feedback from the light emissions of the bulb, which represents an action effect as well. Accordingly, TEC is open to explain a much wider range of phenomena, as we will see below.

**3.1.2. Feature-based coding of perceived and produced events.** In theories on perception, attention, and memory it has become common practice to think of stimuli being represented by composites of feature codes (Allport 1987; 1993). To a considerable degree, this owes to the concentration of research on the visual modality, which again is motivated by the recent progress in understanding the neural basis of visual perception. As we know by now, visual information is projected to several areas throughout occipital, temporal, and parietal lobes, partially following anatomically distinct pathways (e.g., Cowey 1985; DeYoe & Van Essen 1988), and there is no indication that all information belonging to a stimulus or object would converge onto some common "grandmother cell." Instead, different stimulus features coded in different cortical areas seem to be integrated by coordinating the codes representing them. This may be done by modulating and synchronizing the temporal behavior of neural feature codes (for overviews, see Singer 1994; Treisman 1996), but for TEC any other neurophysiological integration mechanism may do just as well.

If we combine the common assumption that stimulus representations are composites of feature codes with our claim that representations of perceptual and action events are of the same kind, an obvious conclusion follows: Action plans should also be made of temporarily composites of action-feature codes. The assumption of a composite action representation is not entirely new. In some sense, the seed for it was already set in the theories of Adams (1971) and Schmidt (1975) on motor learning and in Turvey's (1977) considerations on action control. These authors claimed that the representation of a particular action is not a single, indivisible whole, and not as low-level and muscle-related as Keele's (1968) earlier definition of a motor program might have suggested. Instead, it comprises at least two different parts or structures, such as the perceptual trace and the memory trace in Adams' closed-loop model, or the parameters and invariants in Schmidt's schema theory. Later approaches, as that of Rosenbaum (1980), further devel-


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

oped the idea that response planning involves the specification of action features, and a number of authors have pointed out that these features are coded in distinct functional systems located in different areas of the human brain (Allport 1993; Jeannerod 1997; Keele et al. 1990). That is, actions seem to be represented in a way that is at least very similar to how visual objects are represented. If so, the principles underlying the organization of perceptual and action-related information should be comparable and, in fact, it has been suggested that elements of action plans may be recruited and temporarily bound by similar mechanisms as elements of object representations in perception (Murthy & Fetz 1992; Singer 1994; Stoet & Hommel 1999; see section 3.2.2 on Activation and Integration of Feature Codes).

**3.1.3. Distal coding of event features.** We claim that the cognitive codes that represent perceptual objects are identical to those representing action plans because both kinds of code refer to external, that is, distal events. Importantly, this logic only works if we assume that the respective distal codes represent distal attributes of the perceived event (Heider 1926/1959; 1930/1959; Brunswik 1944) and/or produced event, but not proximal effects on the sensory surface or muscular innervation patterns (cf. Prinz 1992).[3] Consider, for instance, a person reaching for a cup of coffee standing in front of her. One of many possible ways to analyze this situation would be to conceive of the cup as stimulus and of the reaching movement as suitable response. Clearly, a successful response requires that several features of stimulus and action plan match: The intended traveling distance of the hand should be identical with the perceived hand-cup distance, the intended grip should reflect the perceived size of the cup, and the spatial movement goal should be identical with the cup's perceived location. According to our considerations, such a task is easy because stimulus and to-be-performed response share a large number of features. As action planning mainly consists in specifying and integrating the codes representing the intended action features, and as these codes are already activated in the course of perceiving the stimulus, there is not much more to be done (see below). Note, however, that distance, size, and location of stimulus and response only match with regard to a distal description of the environmental layout, but not in terms of the particular neural codes or activation patterns by which it is represented. In fact, there is no way in which the sensory code representing a particular spatial distance would be similar to the muscular innervation pattern driving the hand over the same distance, suggesting that a match or mismatch between stimulus- and action-related codes can only be assumed on a more abstract distal-coding level, and it is this level our approach is referring to.

Distal coding of stimulus objects and action plans has several obvious advantages. First of all, it allows perception and action planning to abstract from domain- and modality-specific (e.g., visual, kinesthetic, or muscular) coding characteristics and refer instead to an event's informational content (for an elaboration on this theme see Prinz 1992). An important implication of the distal-coding notion refers to the number of feature codes available in the coding system and, thus, the grain of possible discriminations in perception and action planing. While the inventory of proximal codes may be of a fixed size, such as with feature detectors in the visual modality, the resolution of distal coding is virtually unlimited. Even though, for instance, the contribution of the auditory modality to frequency perception (i.e., its proximal resolution) is limited, listeners may be able to increase their judgmental abilities by learning to consider other, nonauditory information, such as vibration cues or pain-receptor responses. Similar strategies can help to increase the accuracy and resolution of motor responses, such as when artificial feedback is used to fine-tune a manual response or to gain control over hitherto autonomous functions. Therefore, our claim of feature-based coding should not be taken to mean that the number of feature codes – and, thus, the number of possible discriminations – is given and fixed from birth. Even if this may be so with sensory and motor codes, there is no reason to believe that it is also true for the distal codes TEC is dealing with.

### *3.2. Anatomy and characteristics of event codes*

**3.2.1. Codes of event features.** TEC's core concept is the event code, which again consists of the codes that represent the distal features of an event (or, as a shorthand notation, feature codes). Feature codes are not specific to a particular stimulus or response, but do both register sensory input from various sensory systems and modulate the activities of various motor systems. As feature codes refer to distal event features, they rely on proximal information, but they do not necessarily underlie the same limitations than a given proximal code. For instance, many features of environmental events are available through more than one sensory modality, so that limitations of one modality can be compensated by considering information from another modality, and it is feature codes that integrate this information.

Through integrating information from multiple sources, including memory, distal feature codes can become more complex than proximal codes, which are restricted to a particular sensory channel and reflect the characteristics of particular feature-detection receptors. Therefore, we assume that the dimensions distal feature codes refer to need not always be as simple as color or shape, but can also be as complex as "sit-on-ableness," to take one of the standard "affordances" of Gibsonian approaches. Even time and change might be represented by feature codes, so that events like a leftward motion can be coded.

Also important, feature codes are not simply given but evolve and change through the perceiver/actor's experience. A color will not, or not only and always, be coded as RED, say, but the perceiver/actor will (at least be able to) learn to distinguish several tones of red, such as CRIMSON, ORANGE, and ROSY-RED. As a consequence, a formerly single feature code will become differentiated into a larger number of codes. Likewise, a particular action might not, or not only and always, be coded as LEFT, but one will be able to learn distinguishing, say, LEFT-OF-BODY, LEFT-OF-RIGHT-INDEX-FINGER, and LEFTWARD. In other words, discrimination learning will lead to a continuous change in the ability of a perceiver/actor to represent his or her interactions with the environment. However, we claim that these changes will take place at the level of distally defined feature codes.

A strongly simplified application of our view is given in Figure 1, where two feature codes ($f_1$ and $f_2$) receive input from two sensory systems – say, the visual and the auditory system – and affect the performance of two motor systems – say, hand-movement and speech control.



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

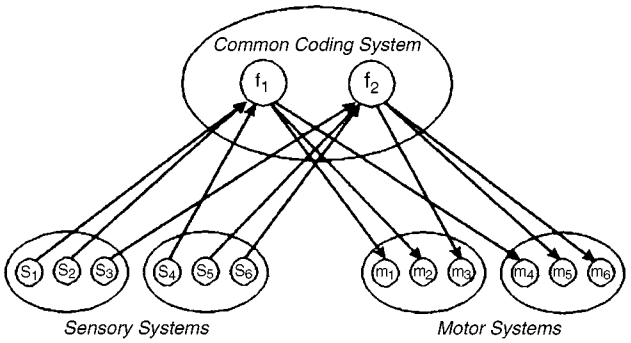



Figure 1. Feature coding according to TEC. In the example, sensory information coming from two different sensory systems ($s_1$, $s_2$, $s_3$, and $s_4$, $s_5$, $s_6$, respectively) converges onto two abstract feature codes ($f_1$ and $f_2$) in a common-coding system, which again spread their activation to codes belonging to two different motor systems ($m_1$, $m_2$, $m_3$, and $m_4$, $m_5$, $m_6$, respectively). Sensory and motor codes refer to proximal information, feature codes in the common-coding system refer to distal information.

Let us assume that, from a perceptual point of view, $f_1$ represents the fact that a particular stimulus tone appears to the left side of the perceiver/actor's body, while $f_2$ represents the, say, high pitch of this tone. Although some spatial information from the auditory system (coded by $s_4$) is used for coding tone location, visual information (coded by $s_1$ and $s_2$) about the (apparent) auditory source may also be considered – a theoretically nonessential attempt of ours to model the well-known phenomenon of visual dominance in spatial perception (e.g., Posner et al. 1976). The perception of pitch is mainly driven by auditory information (coded by $s_5$ and $s_6$), but it may be facilitated by some visual information (coded by $s_3$), such as cues indicating the presence of a violin.

Let us now consider how an action plan is made. Planning an action does not involve specifying every single muscle activity in advance, but is restricted to constraining and modulating sensorimotor coordination to achieve the intended class of movements (Greene 1982; 1988; Turvey 1977). In other words, action control deals with the intended outcome of an action, not with the particularities of the movement or the sensorimotor interplay producing that outcome. According to TEC, an action plan is made of several feature codes, with each code modulating a particular aspect of sensorimotor coordination. To take our example, $f_1$ might control the "leftness" of actions by influencing the activity of motor codes in the hand system (say, $m_1$ and $m_2$) to drive the hand leftward, or toward a left-hand target object. At the same time, the very same code may also bias other motor systems to produce "left" events, such as the speech system ($m_4$) to say "left"[4] or the eye system (not shown) to drive the eyes leftward, thus enabling and supporting multi-effector coordination. Usually, specifying a single feature of an intended action does not suffice, so that action plans will include a number of feature codes. In our example, the code $f_2$ might bias actions to produce "high" outcomes by activating motor code $m_3$ in the hand system to drive the hand upward (e.g., resulting in a hand movement to an upper-left target location) and by affecting parameters $m_5$ and $m_6$ in the speech system to produce a word in a high pitch (e.g., resulting in uttering the word "left" in high voice).

In the proposed common-coding system several kinds of interactions between perceptual and action-related processes are expected, especially if the features of perceived and to-be-produced events overlap. It can be seen from Figure 1 that perceiving an object possessing particular features (e.g., a high-pitch tone) will prime those actions that produce the same features (e.g., speaking in a high voice), and vice versa. Of course, interactions are also expected between processes dealing with different, but feature-overlapping, perceptual events (e.g., a tone and a light on the left side) or actions (e.g., moving the hand and the foot to a left-side target). Even negative effects between different perceptions and different actions or between perception and action are possible, as we will discuss in the next section. Importantly though, these interactions are not due to the characteristics of particular sensory or motor codes, or to some direct interplay between them – the only thing that matters is whether or not they are mediated by the same feature code in the common-coding system. In other words, perceptual and action-planning processes only interact if the codes they operate on refer to the same (kind of) feature of a distal event.

**3.2.2. Activation and integration of feature codes.** Each event code consists of several feature codes representing the attributes of the perceived or planned event. For instance, perceiving a cherry will result in the activation of those feature codes that represent the attributes RED, ROUND, and SMALL, among many others. We have already pointed out that merely activating a particular feature code will prime all those events it shares features with, so that registering the cherry will facilitate perceiving other red, round, and small objects, or performing actions directed towards, manipulating, or producing events possessing these features. This logic works either way, so that selecting the features of a to-be-planned action will facilitate both the perception and the production of other events the planned action shares features with.

However, the mere *activation* of feature codes does not yet make an event code. What if our cherry comes with an apple, which would also be round, but neither red nor that small? Registering the apple should activate the feature codes GREEN, ROUND, and BIG (relatively speaking), so that five feature codes would now be active. How would the system be able to tell, for instance, that the BIG code belongs to the GREEN, not the RED code? Obviously, some kind of *integration* mechanism is required that binds those feature codes together that have been activated by the same event. For the visual domain, several authors have suggested that feature binding is achieved by temporally coupling or synchronizing the activation of feature codes (for overviews see Abeles 1991; Singer 1994; Treisman 1996). Whatever the details of this mechanism may be, an increasing number of studies provides substantial evidence that feature binding in perception does occur. For example, Kahneman et al. (1992) have shown that repeating a particular stimulus is only of advantage if its relative location is also repeated, which strongly suggests that form and location codes of a stimulus object are bound together. Other studies found very similar effects with other stimuli, features, and tasks (Gordon & Irwin 1996; Henderson 1994; Henderson & Anes 1994; Hommel 1998b), which indicates that feature binding is a rather general phenomenon.

Binding problems are not restricted to the perceptual do-


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

main. Assume, for instance, a person is planning to move his/her right hand upwards to pick an apple from a tree and, at the same time, plans a downward movement with the left hand to catch the apple should it fall. To simplify matters, let us assume that only four discriminative codes are involved in specifying these actions: a LEFT and a RIGHT code, an UPWARD and a DOWNWARD code. If action planning consisted of just activating these codes, much confusion would arise in this situation, because it would be impossible to tell whether it is the left or the right hand that needs to be moved upward or downward, thus, whether the LEFT (or RIGHT) code goes with the UPWARD or the DOWNWARD code. This is the same binding problem as discussed for the representation of perceptual objects and it has been suggested that the mechanisms solving it are also the same (Murthy & Fetz 1992; Stoet & Hommel 1999). That is, action plans are not just bundles of activated feature codes, but integrated wholes.

The distinction between activation and integration has profound implications for predicting and understanding interactions between event codes. Figure 2 shows the representations of two events, Event 1 comprising the feature codes $f_1$, $f_2$, and $f_3$, and Event 2 made up of $f_3$, $f_4$, and $f_5$. Note that the two representations overlap in $f_3$, this inviting interaction between them. Now assume that Event 1 is perceived or planned by first registering or selecting (i.e., activating) the corresponding feature codes and then integrating them. As long as the feature codes are merely activated, the representation of Event 2 will also be (partially) activated, this leading to the priming or refreshing of that event code. However, as soon as the features belonging to Event 1 get integrated (indicated by the straight-lined ellipse in Fig. 2), the two event representations no longer support, but interfere with, each other. As $f_3$ is now associated or synchronized with Event 1, it is no longer available for representing other events, so that the integrated code of Event 1 will (partially) suppress or hamper the coding of Event 2 (indicated by the broken-lined ellipse).

This scenario not only allows an account of both facilitation and interference between event-coding processes; it also predicts specific time courses of these phenomena. Regarding perceptual coding, TEC assumes that the first phase of processing a stimulus consists in the parallel activation of all stimulus-feature-related feature codes. If one or more of the activated codes are, or already have been, used to form another event code, performance relying on this event code is facilitated. This is true for perception (i.e., if the respective event code refers to a stimulus event) and for action (i.e., if the code refers to an intended action outcome). Regarding action planning, the first step will also consist in activating the feature codes of the intended action features (a possible exception being highly overlearned actions that may be stored in an already integrated format). Again, this will prime event codes with overlapping features, whether these are used in perception or action planing.

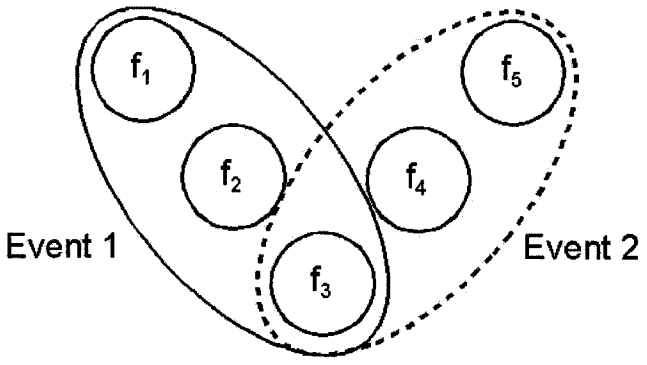



Figure 2. Integration of feature codes into event representations. Feature codes that are activated by external stimulation or internal processes are bound into separate, coherent event structures. In the illustrated example, each of the two represented events includes to two unique features ($f_1$, $f_2$, and $f_4$, $f_5$, respectively), but the two events (and, thus, their representations) overlap with respect to one feature ($f_3$).

During the second phase, activated feature codes will be integrated, so that they are no longer (easily) available for concurrent coding processes. As a consequence, facilitation of processes operating on feature-overlapping events turns into interference. In perception, integration is likely to be associated with attentional processing (Treisman 1988) and, therefore, will depend on whether the respective stimulus is attended by the perceiver (see next section). In action planning, integration is likely to be associated with specific preparation of the particular action, not with the general intention to perform the action on some occasion (Stoet & Hommel 1999).

**3.2.3. Attentional and intentional modulation of event coding.** Perceptual and action-planning processes are selective. For instance, if you take a look at your watch to see what time it is, you might be aware of the watch's orientation relative to you and of the relative location of, and the angle between the hands, because these features are relevant for extracting the correct time information. However, you might not notice – and possibly not remember on a later occasion – the color of the hands or of what material the watch strap is made of, although these features may well have been very salient when you bought the watch. In short, the situational context and the current intentions of the perceiver-actor are likely to have an impact on the processing and coding of perceived events. And the same is true for produced events: If you are reaching out for a cup of coffee, say, it is critical for the success of the action that the fingers know how to grip the cup's handle. Other action features are of much less relevance, such as the hand's exact angle relative to the cup or the parameters of its trajectory, although they are likely to be of central importance if we replaced the cup by, say, a snake. Again, contexts and intentions are crucial in defining what features of an action are relevant and which are not. Moreover, in order to coordinate perceptual and action-planning processes, selectivity on the perception side needs to be tuned to action requirements (selection for action, Allport 1987) and vice versa.

In TEC, event coding in perception and action is highly (although not completely) dependent on the perceiver/actor's current aims and goals. In particular, we assume that event coding is tailored to the situational demands by means of setting and changing the relative weights of feature codes, thereby influencing the degree to which these codes contribute to the resulting event code. If a particular feature is relevant for a given task – whether it codes a stimulus or a response – its code will be primed in advance, which has several consequences. First, the feature code's



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

basic activation level will be increased relative to the standard resting level. Second, if the code then gets activated during the processing of a stimulus event or an action plan, its activation level is higher than that of codes corresponding to task-irrelevant features. For instance, if the form of a stimulus is task-relevant and its color is not, both form and color codes will receive some degree of activation when the stimulus is presented; however, as the form codes are primed due to task relevance, the net activation will be higher in the form than the color code. Third, the higher activation of task-relevant codes entails that these codes will play a dominant part in feature binding: the higher the activation of a code the more strongly – and/or the more likely – it will become integrated into the resulting event code. Consequently, feature weighting affects both activation and integration. If a feature code is weighted highly, its activation level will be higher than that of other, less highly weighted codes – with the result that during the activation phase it will more strongly facilitate the coding of events possessing the respective feature. However, once integration sets in, the respective feature will be more prominently represented in the emerging event code and, therefore, the coding of other feature-overlapping events will be hampered more if they overlap in this than in other, less strongly weighted features.

The feature-weighting principle implies that representations of objects and actions may well include information about both relevant and irrelevant features of the represented entities, yet the former will dominate the latter and, thus, will in a certain sense define or characterize the whole entity. We assume that this is so no matter what kind of entity is represented, that is, features will be weighted according to their task relevance in perception as well as in action planning. With reference to perception, feature weighting may be called an *attentional* process, inasmuch as it selectively prepares the cognitive system for the differential processing of relevant (i.e., to-be-attended) and irrelevant (i.e., to-be-ignored) features of an anticipated perceptual event. With reference to action planning, however, the same kind of feature weighting could rather be called an *intentional* process, because it reflects the perceiver/actor's intention to bring about a selected aspect of the to-be-produced event. In other words, feature weighting always implies preparation for and anticipation of forthcoming events, and this may either refer to a to-be-perceived or a to-be-produced event.

At first sight, the feature-weighting principle might seem to imply that the impact of attentional/intentional processes is restricted to single values on feature dimensions. In fact, however, how the principle is applied is likely to vary with the task. First, there are situations where it is only or mainly a single feature that matters, such as in seeking a star in the sky or when raising a finger. Under these conditions increasing the weight of a single value on a feature dimension (e.g., BRIGHT or UP) will suffice to solve the task. Second, there are situations that require discriminative responses to the dimensional value of a stimulus, such as when being confronted by a traffic light or in a binary-choice experiment. Under these conditions, a first selection of signal stimuli from noise, and of valid from invalid responses, is possible by increasing the weights for a whole feature dimension (Ward 1982), and there is indeed evidence that defining the task-relevant stimulus- and response-feature dimensions is an important part of preparing and implementing a task set (Meiran 2000). Third, given that features can vary in complexity, increasing the weights for a particular feature or feature dimension also implies a selective preparation for a particular level and grain size of event coding. For instance, when being confronted with another person, one can choose to attend to, and act toward the whole person, his/her face, eye, or pupil, and these attentional attitudes are likely to be expressed by weighting codes of event features that differ in complexity. Similarly, in a left-right choice-reaction task, say, actors are likely to specify and code their responses in terms of the categorical feature codes LEFT and RIGHT, while in a positioning task action planning will rely on feature codes of a much finer grain size. That is, perceiver/actor's will switch between more abstract and more detailed representational or coding levels, whatever is more suitable to performing a task.

**3.2.4. Roles of event codes.** Although we deny any fundamental difference between the event codes that underlie perception and those that are functional in action planning, we admit that it still makes sense to distinguish between stimulus codes and response codes. Let us consider a movement with the right index finger, say. This movement can have internal causes, such as when the actor intentionally lifts the finger in response to a signal, or external causes, such as if the finger is mechanically lifted by a motor to signal a response with another effector. Even though the efferent contribution to these two events will grossly differ, the (re-) afferent information will be more or less identical; thus, their codes will largely overlap. Nevertheless, the representations of the afferent information serve different purposes in the two examples: In the first, active situation, the codes referring to the perceived (self-controlled) finger lift represent the intended action goal, and can therefore be considered action or response codes. In contrast, in the passive situation the very same codes serve to represent the (experimenter-controlled) stimulus, and can therefore be legitimately considered stimulus codes. However, we must not forget that it is only the role of the represented event that decides whether a code is a stimulus or a response code, not the features or characteristics of this event or of the codes representing it. That is, what one regards as stimulus and response depends more on who controls the respective event – the experimenter or the participant – and not so much on the nature of its representation. This all comes down to the conclusion that the *role* of an event in a given context should not be confused with the *type* of its cognitive code. Not only can different representations play equivalent roles; the same representation can also play rather different roles.

**3.2.5. Hierarchical coding.** For the sake of simplicity, we have used the term "event" as if it referred to an easily discriminable, well-defined snapshot of the world or a single, discrete movement. Intuitively, it seems justified and not overly problematic to call a light flash or a single keypress an event that is presumably represented by a single, coherent cognitive structure. But what about a series of light flashes or keypresses, or a whole stimulus-response pair? Couldn't they still count as one event? And what about a movie, a holiday trip, or a scientific career?

Obviously, it is rather difficult to imagine how to define an event in a way that is strict and meaningful at the same time. However, this definitional problem does not only ex-


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

ist for the theorist, but for the cognitive system as well. True, designers of good psychological experiments will usually try hard to make it very clear to their participants what counts as relevant stimulus and legal response. Outside the lab, however, it is much less clear how to properly individuate perceptual events and identify their structure. Assume, for instance, you are watching a soccer game or a birthday party. Obviously, such events consist of what Barker (1963) has called a "stream of behavior," which not only could be, but actually is, segmented in tens, hundreds, or thousands of meaningful units, depending on the observer's interest, expertise, or attentional capabilities (e.g., Cohen & Ebbesen 1979; Massad et al. 1979; for an overview, see Stränger & Hommel 1996).

Given this flexibility of the cognitive system to define events internally by selecting a particular temporal interval, a particular aspect, and a particular grain size of the physical event, it would make little sense to come up with an a priori definition, at least at this point of theoretical development. Therefore, we will stick to our loose usage of the term "event" and will leave it to empirical investigation and domain-specific theorizing to lay the groundwork for a stricter definition. Very likely, those efforts will end up with a much more complex picture than we have sketched in Figure 1. Thus, although our examples keep things as simple as possible, this is not to deny the formation of higher-order event codes in the cognitive system. Such higher-order codes may refer to several levels of an event, and they may become themselves integrated into even higher-order representations of whole tasks, and so forth. That is, even if this paper and most of our own research focuses on the lowermost representational level – the relationship and interplay between codes of simple events and codes of their features – this level only provides the basis for presumably much more complex representations of perceived and produced events.

## 4. Empirical evidence

In this section we will mainly review behavioral data from a number of labs including our own covering several fields of research, domains as different as visual attention, action planning, and sensorimotor performance. These studies served different functions in the history and development of TEC – some motivated the conception, inclusion, or addition of theoretical assumptions, some were motivated by TEC and carried out to test its implications, and some coevolved, so to speak, with the theory. All taken together, we think, these behavioral findings are in good agreement with our central assumptions and thus provide ample support for TEC as developed so far. Yet, before going into the behavioral details we briefly mention some recent neuroanatomical and neurophysiological evidence suggesting the existence of brain modules shared by perception and action planning. Obviously, TEC is a purely cognitive approach that is not bound to, or relies on particular brain mechanisms; nor do we wish to claim that the available neuroscientific evidence necessarily requires an approach such as we suggest. Nevertheless, we do find it important to point out that our ideas fit nicely into the current neuroscientific picture.

One example of how perception and action planning may be interfaced in the brain are the "visual-and-motor neurons" found in the monkey's parietal cortex and the "mirror neurons" located in the premotor cortex – areas that are commonly associated with action planning (Jeannerod 1997; Passingham 1993). Visual-and-motor neurons are active when a monkey manipulates a specific object and/or while that object is merely fixated (e.g., Sakata et al. 1995; Taira et al. 1990). Mirror neurons become active when a monkey both performs a given action and observes a similar action performed by an experimenter (e.g., Di Pellegrino et al. 1992; Fadiga et al. 2000; Gallese et al. 1996). Both types of neurons are suggested to play an important role in matching the perceived characteristics of objects to the planning of appropriate movements.

These findings of single-cell recordings point to populations of neurons which seem to fulfill both perceptual and action-planning functions and could be considered as a neuroanatomical substrate of the common codes assumed. Yet, localizing those modules is only one step toward finding a common-coding implementation in the brain; it is also important to identify mechanisms that integrate perceptual and action-related brain activity.

We have already mentioned that one solution to the problem of binding cortically distributed visual activity might be temporary cell synchronization, that is, a mechanism that integrates stimulus features into a coherent whole (oscillations in the gamma range, cf. Eckhorn et al. 1993; Gray et al. 1992). These synchronizations have been assumed to bridge the gap between the visual hemispheres (Engel et al. 1991; Nelson et al. 1992) and, more importantly in the present context, between different motor areas (Conway et al. 1995; Riehle et al. 1997), and between sensory and motor areas of the brain (Roelfsema et al. 1997).

These findings notwithstanding, TEC was primarily guided by the attempt to establish a functional model of perception and action planning applicable to behavioral data. The following overview is organized into four sections. First, we will discuss evidence from object perception studies suggesting that cognitive representations of visual objects are formed by temporarily integrating codes of the perceived event's features, thus, by constructing what Kahneman and Treisman (1984) have called "object files." Second, we present empirical findings showing that action plans share many characteristics with object files, and thus may be formed in more or less the same way. In particular, we argue that action plans, or "action files," are coded in a distributed fashion and in terms of distal features of the action. Third, we show that the principles of object coding and action planning also apply to inter-domain bindings, thus, to the integration of perceptual and action features into common representational structures, and we list evidence for direct and bilateral interactions between object recognition and action planning that exhibit exactly the kind of pattern predicted by TEC. Finally, we discuss findings that support our view that event representations are not invariant and static, but adaptively tailored to the task at hand and the perceiver/actor's intentions.

### 4.1. Coding of perceived events ("object files")

We suggest that representations of objects and other events are constructed by binding separable cognitive codes that correspond to the distal features of the represented entity. During the last decade the idea that coherent central representations may be formed by transiently coupling (ana-



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

tomically or, more important for our approach, functionally) distributed codes has made a considerable impact on psychological theorizing, especially in the domain of visual attention. Treisman and colleagues (Kahneman & Treisman 1984; Treisman & Gelade 1980) were among the first to propose a binding mechanism that serves the integration of distributedly represented visual features. Treisman's feature-integration theory assumes that the registration of a particular stimulus leads to the parallel activation of all those codes that represent the features of that object. However, if more than one object is available, a serial scan of the stimulus array is necessary to glue together, so to say, those subsets of feature codes that correspond to, and thus represent, each given object. That is, visual attention is directed to only one object at a time, thus selectively enhancing the processing of this object's features. A comprehensive feature-based object representation is then established by binding together all these temporarily enhanced feature codes. Only recently, Kahneman et al. (1992) have refined the original feature-integration theory by claiming that feature-code bindings are integrated together with more abstract, semantic knowledge about an object into an "object file." As an integrated, knowledge-enriched structure of specific feature combinations, an object file does not just register features of a perceived object – this could be achieved by merely activating feature codes, not binding them – but also represents an individuated episodic trace of a particular stimulus perceived in a particular context. In other words, it is a true event code.

Empirical evidence for the existence of object files comes from at least two lines of research. First, Allport et al. (1985) reported findings suggesting that the form and the color of a stimulus are integrated into a common representation, so that subsequent presentation of the same form in a different color, or vice versa, leads to a decrement in performance. For instance, they had participants respond to the target stimulus of target-distractor pairs, such as a red target letter superimposed on a green distractor letter. If, on a given trial, the current target matched the preceding distractor (e.g., red A and green B, preceded by red C and green A), reaction times (RTs) were substantially lengthened as compared to trials including non-matches (e.g., red A and green B, preceded by red C and green D). This so-called negative-priming effect has turned out to be quite robust and generalizable over a number of different stimuli, responses, and tasks (for reviews, see Fox 1995; May et al. 1995).

From a feature-binding approach, negative priming might be explained as follows (see Allport et al. 1985; Neill 1997; Park & Kanwisher 1994). Assume that, on a given trial, target features and distractor features are integrated separately, forming two distinct object files (e.g., one containing form and color of the target and one containing form and color of the distractor). If the target is identical with that on the preceding trial, this implies a complete feature match (form and color), thus, a repetition of the stimulus-feature conjunction. If the target does not match the preceding target or distractor, this implies a mismatch, so that new form-color bindings have to be formed. However, if the target matches the preceding distractor – thus, the same form comes up in a new color – this implies a partial feature match. As the task requires the color-cued identification of form, suggesting the integration of form and color information, such a partial match introduces the necessity to undo the old, now invalid and misleading binding before the new one can be formed. Obviously, this undo process takes time, which produces the negative priming phenomenon.

Further evidence for the binding of stimulus features has been reported by Kahneman et al. (1992). They had participants name a single probe letter, which was preceded by a task-irrelevant multi-letter prime display. If the probe letter already appeared as part of the prime display, RTs were slightly faster than if not, thus, there was (however often insignificant) an effect of form repetition. More interesting though, the benefit of letter primes greatly increased when prime and probe shared the same (relative or absolute) location, thus, there were what the authors call "object-specific preview benefits." This suggests that form and location of each prime letter were integrated into separate episodic object representations (i.e., object files). If so, the coding of a probe letter whose form-location combination matched that of a preview element could make use of that element's already formed object file, so that integration was speeded up. Again, this implies that integration is easier with complete than with partial feature matches.

Recent studies have provided ample support for the general notion of stimulus-feature bindings, but they have also helped to refine, modify, and extend the original approach of Kahneman et al. (1992). The first important qualification concerns the question of how object files are accessed. Kahneman and colleagues assumed that, once an object file is established, the only way to access its content is mediated by spatial location. That is, only if a new object appears at about the same location as a preceding one are its features matched against those of the preceding object, so that conjunction-repetition priming as evidenced by negative priming or preview effects should be obtained only with a high degree of spatial correspondence between prime and probe object. However, Hommel (1998b) repeatedly demonstrated effects of spatially unmediated form-color conjunctions, that is, benefits due to the repetition of particular conjunctions of stimulus form and color, independent of whether location was also repeated or not. Such a finding suggests that location information does not (necessarily) play a special role in the access to object files; it rather seems that object files are associative structures that are accessible via any element they include (cf. Leslie et al. 1998).

A second modification addresses the question of how complete these files are. According to the spotlight view of spatial attention as defended by Treisman (1988) and many others, one would expect that attending a particular object leads to the enhanced processing of all the features of this object (the same prediction is derived from the approach of Allport et al. 1985, who maintain that feature integration is a preattentive process). If so, all those features would be integrated and put into the same file, which implies that such a file is a rather comprehensive data structure. However, object-specific preview benefits were demonstrated with letters (Henderson 1994) and words (Gordon & Irwin 1996) even if the letter case was randomly changed from prime to probe presentation. Likewise, Gordon and Irwin (2000) observed object-specific effects with pictures whose left-right orientation changed from prime to probe. Although one might argue that upper- and lower-case versions of letters or mirror-image versions of an object are often very similar, these observations do suggest that not all perceptual stimulus features are integrated into, and main-


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

tained in object files. Why this is so and what the conditions for integration may be will be discussed in some detail below.

A third argument that calls for some theoretical extensions has to do with the distinction between types and tokens and between the activation and integration of feature codes. According to Kahneman et al. (1992), pure repetition priming results from the repeated activation of a given type node, thus, a single, non-integrated feature code in long-term memory; while conjunction-repetition benefits are due to the repeated access to, or the reuse of, the same perceptual object file. An object file is assumed to be a product of the perceptual system, thus, a rather "early" achievement in the processing stream that represents a kind of interface to long-term memory. However, pure feature-priming effects can be demonstrated to show up and to disappear earlier than binding effects (Hommel 2001a): While feature repetition benefits only occur with very short stimulus-onset asynchronies between prime and probe (140–600 msec), benefits of conjunction repetitions set in later and seem to be much more durable (240–3000 msec or longer). This suggests that type activation comes first and is followed by token formation, not vice versa.

In summary, the available evidence strongly suggests that, in accordance with TEC's basic assumptions, (at least visual) objects are cognitively represented by composites of feature codes. These codes are activated by, and thus represent, external stimulus information and are temporarily integrated or bound into coherent representational structures.

### *4.2. Coding of produced events ("action files")*

As already indicated, the original motive of Kahneman and Treisman (1984; Kahneman et al. 1992; Treisman 1992) to introduce the object-file concept and the idea of a temporal coupling of feature codes was to provide a solution to the problem of feature integration and object formation in (visual) perception. However, according to TEC and the common-coding principle underlying it, the logic of cognitive representation and the way cognitive structure is formed is basically the same in perception and action planning. This suggests a rather radical proposal: Not only should object and event perception be mediated by object-file type of structures, but action planning should be mediated by the same entities as well. We have already pointed out that information about action can only be acquired by means of perception, thus, is always perceptual information – possibly enriched by long-term knowledge. Therefore, the representation of an action should not be very different from the representation of any other event, only that the main function of the former is to constrain the motor system to produce the action-contingent perceptual effects the action aims at. If so, action plans should be very similar to other event representations, such as object files. In fact, a number of studies have provided abundant evidence that actions are coded and planned in terms of features referring to their distal, perceptual effects, suggesting the existence of what one may call "action files."

**4.2.1. Feature-based representation of action plans.** The general assumption that action plans are made up of parts is well supported by empirical evidence. For instance, in Rosenbaum's (1980) classical study subjects performed speeded hand movements with the left versus right hand, for a short versus long distance, toward versus away from their body. Precues appearing before the actual stimulus informed about some or all response parameters, so that subjects had the opportunity to partially or completely make up their action plan in advance. Indeed, initiation times decreased with an increasing number of precued parameters, suggesting that subjects were able to specify the features of their action relatively independently. Later studies extended these findings to other action features and effectors, and to multi-step actions (for an overview, see Rosenbaum 1987), showing that feature-based action planning is a general phenomenon. Further evidence comes from studies focusing on Schmidt's (1975) distinction between programs and parameters. For instance, Quinn and Sherwood (1983) observed that changing the speed of a movement takes much less time than changing direction. Like the findings of Rosenbaum (1980), this suggests that action plans consist of separable pieces the modification of which takes different amounts of time (for an overview, see Schmidt 1988).

TEC does not only assume that action plans are made up of feature codes, it also claims that these codes are integrated, that is, temporarily "glued" together, much as was demonstrated in the studies on object-file formation in visual perception. These studies revealed that the feature codes belonging to one object are integrated in a way that has an impact on subsequent access to these codes. Specifically, if presenting an object results in the binding of two given feature codes (say, ROUND and RED), it is easy to code another, later-presented object if it has the same feature combination (ROUND and RED) or if its features do not overlap at all (RECTANGULAR and GREEN), but it is difficult to code an object with only partially overlapping features (ROUND and GREEN). According to our interpretation, this is because feature integration leads to an associative connection between the respective feature codes, so that activating one element of the emerging structure will spread activation to the other elements, which is helpful in case of complete overlap and may be harmless with no overlap, but is distracting and interfering in cases of partial overlap. In other words, reusing an event code or forming two unrelated ones is not problematic, but forming two overlapping codes is. If action plans are like object files and thus are formed by the temporary binding of action-feature codes, one would similar effects. In particular, if the distinctive features of two responses $R_1$ and $R_2$ overlap, it should be difficult to plan $R_2$, because the critical feature code would be already bound to the plan of $R_1$ and therefore not (or less easily) be available for $R_2$. As a consequence, response initiation for $R_2$ should be delayed in case of $R_1$-$R_2$ feature overlap.

First evidence for problems associated with partial feature overlap of action plans comes studies on speech planning by Meyer and Gordon (1985) and Yaniv et al. (1990). In these studies subjects first prepared the speeded verbal production of one of two possible syllables (e.g., "up"). Then they were signaled either to actually produce the prepared utterance or to utter the alternative syllable, the interesting measure being the additional time needed to replan an utterance as a function of the similarity between the endings of prepared and eventual syllable. As predicted from TEC, partial overlap of these endings (e.g., with respect to voicing, like "up"-"ut" or "ub"-"ud") produced *longer* re-planning times than no overlap (e.g., "up"-"ud" or "ut"-"ub").



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

Further evidence for aftereffects of feature binding in action planning might be taken from a study of Rosenbaum et al. (1986). These authors had subjects verbally repeat the first 2–9 letters of the alphabet with alternating stress (e.g., AbCdAbCd . . . ), which led to much better performance with even-numbered than odd-numbered sequences (e.g., AbCaBc . . . ). From a TEC perspective this might be due to that even-numbered sequences leave the bindings of letters and stress parameter constant, whereas odd-numbered sequences require new bindings in every trial, which is likely to invite interference from the previous binding.

Stoet and Hommel (1999; Stoet 1998) investigated the effects of feature overlap by using a dual-task paradigm with two stimuli ($S_1$ and $S_2$) and two responses ($R_1$ and $R_2$). First, participants saw $S_1$ that signaled a multi-step response with the left or right hand. However, $R_1$ was only planned and prepared, but not yet executed. After some time, $S_2$ came up and was immediately responded to by a speeded left or right $R_2$; only then the already prepared $R_1$ was performed. The crucial question was whether partial feature overlap between $R_1$ and $R_2$ would delay the initiation of $R_2$. Indeed, this is what was observed, even if $R_1$ and $R_2$ were performed with feet and hands, respectively, on the same versus different sides of the body. This shows that, first, overlap costs are not bound to reusing a particular effector and, second, that the overlapping features can be fairly abstract. Moreover, overlap costs turned into benefits if the time interval between $S_1$ and $S_2$ was reduced and participants were discouraged to plan $R_1$ ahead of $R_2$ – a demonstration of both the activation and the integration phase of structure formation within the same experiment.

A further finding of Stoet and Hommel (1999) may point to interesting differences in the temporal characteristics of object and action files. Recall that the studies on object-file formation indicated that forming a particular object file affects the subsequent formation of other object files, which strongly suggests that the lifetime of a given file outlasts the duration of its actual use. Applied to action planning, this implies that once an action plan is formed, it might affect the formation of other, feature-overlapping plans not only during the time it is held in preparation, but even after it has been used to execute the respective action. In the design of Stoet and Hommel, this would mean that $R_1$, which was carried out immediately after $R_2$, should have been delayed with feature overlap between $R_1$ and $R_2$. This, however, was not the case; overlap rather tended to facilitate the performance of $R_1$, especially if the interval between $R_1$ and $R_2$ was short. According to the feature-binding logic, this does not necessarily mean that the processes of feature activation and integration are any different in perception and action planning, but it does imply that the aftereffects of binding differ: Unlike object-related event codes, activation of action-related codes seem to dissolve immediately after use. Although a systematic investigation of this theoretically important hypothesis has not been undertaken yet, we will see some more support for it below, when we will discuss effects of action on perception and interactions between perceptual and action-related processes.

**4.2.2. Distal coding of action.** Up to now we have dealt with spatio-anatomical characteristics of hands and feet, thus, with "intrinsic" action features. One of the first empirical indications that actions can also be coded in terms of distal, "extrinsic" action-contingent events was observed rather accidentally by Morin and Grant (1955). In their study, participants faced horizontal rows of red stimulus lights and, directly below these, green response-information lights. In each trial, two stimulus lights were flashed and were responded to by pressing two of the response keys that were mounted in parallel to the lights. In most conditions the mapping of stimulus lights and response keys was not compatible and rather complicated, so that considerable practice was necessary to learn the correct mapping. To facilitate learning, pressing a key illuminated the green response-information light below the assigned stimulus, so that the task was to switch on the green lights below the red stimuli. After a fair amount of practice, which brought RTs to an asymptote, keys and green information lights were disconnected. Interestingly, performance immediately declined to prepractice level, thus, withholding the action effects eliminated any learning benefits. This shows that learning must have mainly consisted in the acquisition of associations between stimulus-light and information-light positions, that is, in forming linkages between stimuli and action-contingent effects. Of course, this would be of little help if the light-producing actions were not themselves associated with the codes of their visual effects, suggesting that participants were indeed able to code their keypresses in terms of the corresponding light flashes.

There is further evidence for the distal coding of action from studies on spatial S-R compatibility. Typically, those studies require spatial responses to spatially corresponding and noncorresponding stimuli, with the standard outcome that correspondence allows for better performance than noncorrespondence. Correspondence effects are commonly attributed to processes of stimulus and response coding, the assumption being that similar codes are easier to translate into one another than dissimilar ones. However, where these codes come from has not been a major theoretical issue so far.

A first step toward a better understanding of spatial action coding has been made by Simon et al. (1970) and Wallace (1971; 1972). These authors had participants perform standard spatial S-R compatibility tasks, sometimes with hands held in parallel, as usual, and sometimes with hands crossed, so that now the left hand operated the right key and vice versa. Surprisingly, this manipulation had not much of an effect: Correspondence between stimulus and key was associated with better performance than noncorrespondence under either hand condition, and the size of this effect was more or less the same. That is, the keypressing actions in these studies were obviously not coded in terms of the operative effector, but with regard to the location where it performed. This fits well with the results of Heister et al. (1986; 1987), whose participants operated left and right response keys with fingers of the same hand. Not only were left-key responses faster to left stimuli and right-key responses faster to right stimuli – which confirms that spatial compatibility effects are not restricted to the use of left versus right hands – but this applied as well when the hands were held palm up or palm down, which is strong evidence for the effector independence of action coding. And even hand or finger location is not always of relevance: In an experiment of Riggio et al. (1986), participants operated left and right response keys with sticks that were held in parallel or crossed. As it turned out, responses were faster with spatial correspondence between stimulus and response key (i.e., stick-end location), irrespective of the performing hand and its location.


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

One commonality among the studies concerned with dissociating the impact of effector, effector location, and response key is that effector and response event were always mechanically connected. That this is not essential is nicely shown by the findings of Guiard (1983) and Michaels and Stins (1997). In both studies, participants rotated a steering wheel in a right or left direction (i.e., clockwise or counterclockwise), as signaled by the pitch of a stimulus tone. Just as with key responses, leftward turns were fastest if signaled by a left-side tone and rightward turns were fastest with a right-side tone, independent of where and how the hands were positioned on the wheel or in which direction they were to be moved. Obviously, it was the correspondence between stimulus and intended wheel rotation that determined the effect's direction, which provides additional support for the assumption that actions are coded in terms of their distal features.

**4.2.3. Static and dynamic action effects.** Up to now we have only considered paradigms and experimental tasks using static stimuli and static responses, that is, events occurring at fixed positions in space. Although static events are much easier to realize in an experimental situation than dynamic events and are often much simpler to analyze, it should be clear that the framework provided by TEC is not at all limited to these kinds of stimuli, responses, or paradigms. In fact, a considerable number of studies with dynamic events have been undertaken by other researchers and in our lab, with outcomes that are in good agreement with expectations from TEC.

One of the first studies to show that effects of S-R compatibility can not only be obtained with static stimuli, but with moving objects as well, was conducted by Michaels (1988). In this study, participants were presented with two rectangular objects at the same time, one on the left and one on the right side of a screen. After a delay one object started moving to the left or right, with the starting position of this object signaling the response. Interestingly, responses were quicker if their location corresponded to the direction in which the object moved, which shows that the task-irrelevant feature of movement direction could not be ignored. That is, dynamic properties of a stimulus event are automatically coded and they activate feature-overlapping responses just as static stimuli do, an observation that could be replicated several times (Ehrenstein 1994; Nattkemper & Prinz 2001; Proctor et al. 1993).

Only recently, we were able to show that very similar compatibility effects can be obtained with even more complex stimulus and response patterns, such as hand postures. In the study by Stürmer (1997; Stürmer et al. 2000), participants faced a color monitor on which postures of a human model's right hand were presented, such as the end states of a grasping or finger-spreading movement. At some point in time, the color of the model's hand changed from skin color to either red or blue, thus signaling to the participant which response to make. The signaled response could be either a grasping movement or a finger-spreading movement (starting from a neutral, hand half-open posture), that is, a movement resulting in a posture that was virtually identical to one of the postures presented on the monitor. As color, not posture, was the relevant stimulus feature, there were compatible and incompatible relationships between the visual display and the required response: Sometimes the response corresponded to the displayed posture and sometimes it did not, this creating a kind of posture-based Simon task. As expected, the time it took to initiate a movement response strongly depended on whether its goal posture did or did not match the displayed posture: Correspondence was associated with quicker responses than noncorrespondence, especially if the color change coincided with the onset of the posture presentation.

Apart from demonstrating that Simon-like compatibility effects can be observed with rather complex stimulus and response patterns, these findings speak to the role of action-goal representations in response selection. Obviously, it was an image of the end state of a movement that facilitated or interfered with selecting the correct response, which clearly suggests that end-state or goal representations play a crucial role in response selection, just as emphasized by TEC. It is true that one may object to bringing goals into play by pointing to the relative similarity of displayed postures and the first phase of the response movement. For instance, if a participant is signaled to spread his/her fingers apart, the image of a wide-opened hand does not only correspond to the intended movement goal, but it is also more similar to the very beginning of the response, the opening phase and its direction, than an image of a closed hand. To test that, further experiments of Stürmer and colleagues compared the effect of goal-posture images with those of pictures showing intermediate movement states. If it were the similarity between displayed posture and movement beginning or direction that produced the effect, one would expect the compatibility effect to get stronger the more similar the displayed posture is to the first phase of the movement, thus, the earlier the depicted movement phase. However, the results showed the opposite pattern, with the strongest effects associated with goal-posture images.

Having been successful in demonstrating compatibility effects with static hand postures and dynamic responses, Stürmer and colleagues went on to extend their paradigm to investigating dynamic stimuli as well. To this aim, they presented to their participants short video-clips showing the right hand of a human model. Two movements could be presented, to both beginning and ending in a neutral half-open position, one showing the hand closing and then opening again (grasping), and one showing the hand opening and then closing (spreading). As before, the displayed gesture did not provide any relevant information and was to be ignored, while the required grasping or spreading response was signaled by a color change of the monitor. Nevertheless, compatibility between model movement and response movement had a strong impact on RT. When the color change occurred during the first half of the displayed movement (i.e., during the closing phase of the grasping movement or during the opening phase of the spreading movement), the initiation of movement-compatible responses (i.e., grasping and spreading responses, respectively) was facilitated. However, when the color change occurred during the second phase (i.e., during the opening phase of the grasping movement or during the closing phase of the spreading movement), responses were quicker if they mimicked that part of the model movement that was currently observed. That is, grasping responses were faster with the (grasping-like) closing phase of the spreading movement on the monitor and spreading movements were faster in presence of the (spreading-like) opening phase of the grasping movement. Obviously, the very same movement can be facilitated by perceiving a representation of its intended goal as well as



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

by perceiving the spatio-temporal movement itself. This provides further support for our assumption that – although the intended action effect plays a special role in action coding and response selection – cognitive action representations are not confined to goal codes, but include codes of other, less crucial action effects as well.

However, note that the findings of Stürmer and colleagues are not fully conclusive as to the question whether goal- and movement-related action effects are cognitively represented at the same time. Assume, for instance, that for some reason participants tend to code their responses in terms of end postures when posture images are used as distractor stimuli, but prefer movement-related action codings with movement distractors. If so, posture images would affect response selection in the first case and movement displays in the second, without requiring end postures and movements being coded in the same task. In other words, people may code their actions in terms of static goals *or* in terms of dynamic movement features, but not both. Although this is a possibility, there are some reasons to believe that both static and dynamic action features are concurrently represented. A first indication comes from studies by Michaels (1988) and Ehrenstein (1994), where participants responded with spatial responses to the direction in which the visual stimuli were moving. Although the stimulus' starting location was not relevant for this task, its spatial correspondence or noncorrespondence with the response strongly affected RT, which at least suggests that starting location, a static feature, and movement direction, a dynamic feature, were both coded for stimuli and responses.

Further indications for concurrent coding of static and dynamic features is provided by Nattkemper and Prinz (2001). In this study, participants monitored a white dot on a computer screen that moved back and forth on a horizontal path. At some point in time, the moving dot's color changed to red or blue to signal a left- or rightward movement with a hand-held stylus. Although this time neither the direction of the dot nor the location of color change was relevant to the task, both had an impact on the speed of response initiation: Leftward responses were faster if the color change occurred on the left rather than on the right side of the screen and when the change occurred while the dot was traveling in leftward rather than in rightward direction, while the opposite was true for rightward responses. Obviously, both the dot's current location and motion direction were coded, and either code had access to and impact on response-selection processes, which nicely demonstrates that dynamic stimulus-response features do not behave any differently from the much better investigated static features.

**4.2.4. Natural and artificial action effects.** TEC assumes that actions are coded in terms of their perceivable effects, thus action-contingent events. This assumption does not leave much space for the distinction between intrinsic, natural, body-related action effects on the one hand and extrinsic, artificial feedback on the other, a distinction that has been emphasized by Salmoni et al. (1984) and others. In fact, TEC leads one to doubt that such distinctions make sense at all, because for a cognitive system any information is necessarily extrinsic and natural – the only thing that counts is the event the information refers to. If this is so, and if actions are coded in terms of their perceivable effects, it should be possible to show that actions can be coded and planned in terms of new, conditioned events that are contingent on the action. Although we have only begun to investigate this implication, there is already some empirical evidence in support of it.

A very first hint to a crucial role of newly learned, artificial action feedback in action planning can be taken from the already mentioned study by Morin and Grant (1955). The observation that the benefits of learning were eliminated completely as soon as the keypressing actions were no longer followed by the flashing of the assigned response-information lights shows that the participants of this study must have acquired some knowledge about the relationship between actions and visual effects, which then was used for action control.

The same conclusion can be drawn from some accidental observations of Guiard (1983) in Experiment 3 of his wheel-rotation study. In this experiment, all participants occupied the same starting position and held the steering wheel bimanually at its bottom (6:30) position, so that a left-hand turn required a rightward movement of the hands and a right-hand turn a leftward movement. One group received no extra feedback (in addition to the intrinsic sensory information provided by moving hands and wheel), while another group saw a light wandering to the left or right, depending on the direction in which the wheel was rotated. The feedback group exhibited a strong effect of spatial compatibility between the stimulus and the direction of the feedback light, but the results of the no-feedback group were mixed. About one half of the participants showed the same effect as the feedback group, while the other half produced the opposite effect: better performance if a rightward (clockwise) turn was signaled by a left stimulus or a leftward (counter-clockwise) turn by a right stimulus, as compared to the correspondence conditions. Given the positions occupied by the hands, stimulus-rotation correspondence meant noncorrespondence between stimulus and hand movement, and vice versa, which implies that the latter half of the no-feedback group coded their actions in reference to the direction of hand movement, not wheel rotation. In other words, if there are alternative, equally salient action effects, such as perceived wheel rotation and perceived hand movement, people are not consistent in choosing one or the other for coding their actions. However, if a situation affords a salient action effect, such as a moving light in a darkened room, this "extrinsic" event is (or at least can be) preferred over "intrinsic" alternatives. Although this is a post-hoc interpretation, it should be noted that the results of Guiard have been replicated only recently by Michaels and Stins (1997).

To overcome the weakness of post-hoc explanations, Hommel (1993) made an attempt to directly control the preferences of participants for one or another available action effect by manipulating the instruction given to them. In his Experiment 1 people responded with a left- or right-hand keypress to the pitch of a tone that appeared randomly on the left or right side. Importantly, pressing one of the keys flashed a light on the opposite side, so that at least two alternative sources of action-contingent stimulation (hand-/key-related vs. light-related) were available for action coding. One group was instructed in terms of keypressing ("press the left-right key in response to the low-high pitch"), while the other group received a light-related instruction ("flash the right-left light in response to the low-high pitch"). As expected, the key group performed better


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

under stimulus-key correspondence (i.e., stimulus-light noncorrespondence) than under noncorrespondence (i.e., stimulus-light correspondence), while the opposite was true for the light group. Obviously, the instruction had the expected effect in persuading the key group to code their actions in terms of proximal effects (keypressing or finger moving) and the light group in terms of more distal effects (light flashing). It is interesting to note that, if anything, the light group was somewhat quicker overall than the key group, which suggests that coding actions in terms of artificial, "extrinsic" effects is no less easy and natural than referring to more "intrinsic" feedback sources.

Further studies have shown that people do not need to be explicitly instructed to integrate new, artificial feedback into their action representations. In a study on the role of task-irrelevant action-contingent events, Hommel (1996a) asked participants to perform very simple binary-choice tasks, such as pressing a key once or twice in response to a red or green color patch. Each response had a particular auditory effect, which could be a low-pitched tone following a single keypress and a high-pitched tone following a double press. After an acquisition phase, each visual reaction stimulus was accompanied by an irrelevant prime, a tone of the same pitch as one of the auditory action effects. If, as TEC proposes, the tone effects were integrated into the event codes of the actions they follow, these primes should have specific effects: They should activate the corresponding action's representation and thus speed up its retrieval and/or interfere with the retrieval of other, alternative actions. If so, performance should be better if a prime matches the to-be-expected effect of the correct response as compared to a mismatch (i.e., a match of prime and incorrect response). This is exactly what was found in several tasks using different stimuli and responses, that is, prime-effect correspondence produced better performance than noncorrespondence. Although these effects are often not very impressive in size, they are rather robust, generalizable, and stable over time. For instance, Elsner (2000) observed reliable effects even after five sessions of about 380 trials each. Among other things, this rules out curiosity-based interpretations of action-effect integration: People should not be too excited any more about two sinus tones after 1,900 or so trials.

More support for the idea that action-related event codes are made up of representations of their perceivable effects – so that actions are controlled by the anticipation of their effects – comes from studies on sensorimotor synchronization. In synchronization tasks participants are asked to tap with a finger in synchrony with a periodical sequence of auditory clicks, that is, to time their actions so as to coincide with certain events. It is commonly observed that people are not very exact in synchronizing; typically, the tap leads over the click by about 20 to 50 msec (e.g., Aschersleben & Prinz 1995; 1997; Fraisse 1980; see Aschersleben, in press, for an overview).

Recent TEC-inspired accounts of this so-called "negative asynchrony" or "anticipatory error" have focused on the way people might cognitively control their performance in synchronization tasks. In particular, following Paillard (1949) and Fraisse (1980) some of us have suggested that synchrony is not only controlled by, but also established at, a central representational level, where according to TEC both stimuli and actions are represented in terms of their sensory effects (Aschersleben & Prinz 1995; 1997; Aschersleben et al., in press; Gehrke 1996). Therefore, action control in synchronization tasks is not so much concerned with the temporal relationship between the auditory input and the motor output but with realizing synchrony between perceived click and perceived tap. If so, the time it takes to perceive the click and the tap becomes crucial, the more so as auditory and kinesthetic-tactile stimulation can be assumed to differ in processing speed. In fact, as the temporal delay between actual and perceived click is likely to be shorter than the delay between actual and perceived tap, the actual tap must precede the actual click to achieve synchrony between the perceived events at a central level, thus, the negative asynchrony between click onset and overt tap is observed.[5]

How can we know that the tap and its temporal characteristics are really cognitively represented by the tap's sensory feedback and not, for example, by the motor command producing it? An obvious way to test this assumption is to manipulate the feedback coming from the tap or, more precisely, the temporal delay between actual and perceived tap. The longer this delay is, the longer the actual tap must lead over the actual click in order to achieve coincidence of the two perceived events, the more pronounced the negative asynchrony should be. In fact, this prediction has found empirical support through several experimental manipulations.

First, participants were asked to tap with different effectors, such as hand or foot, the assumption being that the temporal delay between actual and perceived tap increases with the "neural distance" between effector and brain. As expected, the anticipatory error was more pronounced with foot than with hand tapping, and this effect was independent of the body side involved and of whether one or two effectors performed the movement at the same time (Aschersleben & Prinz 1995; Aschersleben et al., in press).

Second, the tap's natural action effects were manipulated by having people produce taps with strong and weak forces. Forceful taps are characterized by a stronger force applied at the fingertip and a higher velocity of the downward movement, thus yielding a higher intensity of the reafference compared with soft taps. Such an intensification of the sensory feedback should reduce the delay between actual and perceived tap (due to faster accumulation of sensory information at a central level) and thus reduce the anticipatory error. In fact, this pattern of results was not only observed for the standard tapping task, but for isometric force pulses as well (Aschersleben et al., in press; Gehrke 1995; 1996).

Third, and this is the most direct test of the action-effect assumption, we investigated the effects of eliminating (part of) the action's feedback signals. For the tactile component of the tap's somatosensory feedback this can be achieved by means of local anaesthesia applied to the tapping finger, a method that inhibits the transmission of afferent signals from the tactile units of the skin to the brain without affecting kinesthetic and – more important – efferent pathways. If people perform under these conditions, the anticipatory error substantially increases in the standard tapping task as well as with isometric force pulses (Aschersleben et al. 2001), which provides direct evidence for the importance of natural action effects in the temporal control of actions. This notion has gained recent support in a study with a deafferented patient being able to tap in exact synchrony under conditions with auditory pacing signal



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

and auditory feedback (but no tactile/kinesthetic feedback; Aschersleben et al. 2002).

We have already emphasized that according to TEC actions can be coded and planned in terms of "natural" as well as "artificial" conditioned action effects. As TEC does not even assume any qualitative difference between these two types of action effects, they should both become integrated with the action they follow, so that activating any one effect code leads to the activation of the associated action. For the sensorimotor synchronization task these assumptions lead to the following predictions. By presenting after each tap an artificial action effect, such as a tone, one should be able to manipulate the amount of the anticipatory error. This is because if the action effect is actually integrated and becomes part of the tap's cognitive representation, the tap has at least two effects, a somatosensory one and an auditory one. If the perceived time point of the tap takes into account the temporal characteristics of all of its perceived effects, and if, thus, somatosensory and auditory effects are "temporally averaged," this implies that presenting the auditory effect simultaneously with the tap makes the whole tap to be perceived to come earlier. This is because the temporal delay between the actual feedback tone and the perceived feedback tone is shorter than the corresponding delay for the tap. To achieve synchrony between perceived click and perceived tap, the actual tap has to appear later than in the absence of the auditory action effect, thus leading to a reduction in the anticipatory error. Moreover, if the somatosensory and auditory effects are temporally averaged, the reduction in the asynchrony should depend on the size of the asynchrony under "natural" feedback conditions. For example, in the foot tapping task the reduction in the asynchrony resulting from an auditory feedback tone should be greater than in the hand tapping task. In fact, this pattern of results is exactly what has been demonstrated in corresponding experiments (Aschersleben & Prinz 1995; 1997; Mates & Aschersleben 2000; Mates et al. 1992).

The idea of an temporal averaging process has been tested in an even more direct way by varying the temporal interval between the tap and the feedback tone. The introduction of a delay between the tap and the feedback tone should lead to a change in the timing of the tap. Indeed, the more the feedback tone is delayed the earlier the tap is performed and the larger the anticipatory error (Aschersleben & Prinz 1997; Mates & Aschersleben 2000). Meanwhile, the basic idea underlying the temporally averaging assumption has been also successfully applied to the impact of knowledge of results (Aschersleben 2001).

### *4.3. Interactions between action planning and object perception*

We have now discussed in some detail how codes of stimulus features are integrated to form some kind of stimulus representation, and we have reported evidence suggesting that forming an action plan follows very much the same principles. Such an equivalence of processes and principles clearly favors the idea embodied by TEC that the coding of perceived and produced events takes place in the same representational system, probably using the same (kind of) cognitive codes. However, mere equivalence does not really prove the common-coding idea because, after all, it may well be that the same principles are followed in different and unrelated representational domains. Therefore, our next step is to show the existence of cross-domain integration and of between-domain interactions, that is, of binding effects involving both stimulus and action features and of bilateral interactions between perception and action planning.

**4.3.1. Integration of perceived and produced events.** First evidence for the binding of stimulus and response features comes from a study by Hommel (1998b). In this study, participants performed two reaction-time tasks in a row, one requiring a simple, prepared response, the other a binary-choice reaction. In each trial, they were first informed whether the first response ($R_1$) should be a left- or right-hand keypress, then they prepared the response, and waited for a trigger stimulus ($S_1$) to perform it. One second later, another stimulus appeared ($S_2$), whose form signaled another left or right keypress to be carried out immediately. $S_1$ and $S_2$ varied in form, color, and (top or bottom) location, but the only task-relevant feature was the form of $S_2$. One of the main questions in this study was whether response-repetition effects could be observed in $R_2$. Still more interesting was, whether these effects would depend on the repetition or alternation of stimulus features, thus whether $R_1$-$R_2$ repetition interact with the relationship between $S_1$ and $S_2$.

Indeed, there were two interactions of this sort. One concerned the repetition of stimulus form and response: Repeating a response was beneficial only when stimulus form was also repeated, while response alternation produced better performance with the form alternation than with form repetition. This shows that $R_1$ must have been associated with the form of $S_1$, so that repeating the same or an unrelated, nonoverlapping form-response conjunction produced benefits, while only partial repetitions yielded costs. A follow-up experiment with color as relevant $S_2$ feature produced very similar results, only that now it was the color of $S_1$ that was bound to $R_1$, while form played a minor role. Obviously, the integration effect does not depend on a particular feature dimension, but is sensitive to the task relevance of the features involved, an issue that we will discuss in more detail later.

A second interaction effect was related to response and stimulus location, and its pattern was very similar to the first one: Response repetition yielded benefits with the repetition of stimulus location, but a cost with stimulus-location alternation, which is further evidence for the common integration of $S_1$ and $R_1$ features. Yet unpublished control experiments showed that this effect is also dependent on task relevance: If the left-right response in $R_1$ and $R_2$ is replaced by a spatially neutral response, such as pressing a key once or twice, response repetition effects are no longer modified by the spatial relationship between $S_1$ and $S_2$.

A further study (Hommel 2001b) investigated the temporal dynamics of stimulus response binding, and especially the question of whether object files and stimulus response binding exhibit comparable time courses over SOAs of 600–3,000 msec between $S_1$ and $S_2$. Stimulus response binding emerged under all conditions, from the shortest SOA on to the longest, and thus showed exactly the same temporal characteristics as object files, that is, as within-domain bindings between stimulus features. Taken together, these observations provide further evidence for the assumption of common mechanisms and shared structures of feature binding in perception and action planning.


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

**4.3.2. Action effects upon perception.** TEC does not only imply that codes can be integrated across different functional domains, it also predicts specific effects of code integration in one domain upon processes dealing with codes in another domain. One such example concerns the effects of action planning upon the perception of feature-overlapping stimulus events. The observation that performing a manual task can lead to unspecific impairments of a concurrent perceptual task is not new (De Jong 1993; De Jong & Sweet 1994; McCann & Johnston 1992; Ruthruff et al. 1995). For example, De Jong (1993) showed that the identification of a single, foveally presented letter is postponed when the observer is engaged in responding to a presented auditory stimulus (see, also, De Jong & Sweet 1994). Although such a finding provides some evidence for capacity sharing between, or a structural bottleneck common to perceptual and action-related processes, it does not necessarily require the common-coding assumption. However, in a series of recent studies we were able to demonstrate that interference from action upon perception is more specific than a capacity or bottleneck view would suggest.

The basic procedure we used includes a dual task, in which participants see a tachistoscopically presented, briefly masked arrow while performing an already prepared, unspeeded left or right keypress. After the keypress, participants judge the direction of the arrow, which randomly points to the left or right side. According to TEC, planning and performing a left- or right-hand keypress requires integrating a LEFT or RIGHT code, respectively, into the corresponding action plan. If so, this integrated code should be not (or at least less) available for processing and coding a LEFT or RIGHT arrow, so that people would be effectively "blind" to arrows that point to the same side as the response. As a consequence, perceptual judgments of arrow direction should be impaired if, say, left-pointing arrows appear during left-hand keypressing or if right-pointing arrows appear during right-hand responding. In other words, feature overlap between action plan and stimulus should impair stimulus processing.[6]

Indeed, this is what we observed in a number of experiments. We also found that the perceptual arrow-judgment task is much easier when the arrow is presented without a concurrent task than when it appears during the planning or the execution of the left-right keypress (Wühr & Müsseler 2001). This result is in agreement with previous demonstrations of unspecific interference from response generation upon perceptual identification (De Jong 1993; De Jong & Sweet 1994; McCann & Johnston 1992; Ruthruff et al. 1995). More important from a TEC point of view, however, was the observation of specific impairments with feature overlap between response and arrow direction: If the arrow pointed to the side of the response, it was about 10% less accurately reported than if arrow direction and concurrent response did not match (Müsseler & Hommel 1997a).

This basic finding was corroborated in several series of experiments that aimed at ruling out alternative interpretations. For instance, we found the specific impairment to be independent of the relationship or similarity between the arrow and the stimulus by which the keypress was signaled (e.g., a word, a tone, another arrow, or no stimulus at all, Müsseler & Hommel 1997a; Müsseler et al. 2000), or of the relationship between the keypressing response and the response given during the arrow-judgment phase (e.g., a verbal response, a mouse click, or another keypress; Müsseler & Hommel 1997a). Further, we were able to show that feature overlap produces "blindness" even in detection tasks independent of whether the impairment was measured in terms of percentage correct (Müsseler & Hommel 1997b) or in terms of the signal-detection parameter *d'* (Müsseler et al. 2001). That is, the effect is not bound to stimulus identification tasks. Other experiments extensively manipulated the temporal relationship between the manual response and the presentation of the arrow. As expected from TEC, arrow blindness is not restricted to the execution phase of the response, but also occurs while the response is being planned and held in preparation (Wühr & Müsseler 2001). Interestingly, no effect was observed if the arrow appeared immediately after response execution, which suggests that the integration of LEFT and RIGHT codes into action plans does not survive the functional lifetime of the corresponding plan.

Further evidence for interactions between event perception and action planning comes from other types of overlapping tasks (Schubö 1998; Schubö et al. 2000; see also Müsseler & Prinz 1996). In the task introduced by Schubö on each trial n, participants watched a sinusoidal motion on a computer ($S_n$), while attempting to copy the motion they saw in the preceding trial ($S_{n-1} \rightarrow R_n$). According to the framework provided by TEC, a task like this is extremely interesting to study because it implies two conflicting requirements: to both process the two events ($S_n$ and $R_n$) simultaneously (as they co-occur in time) and still keep them separate (as they belong to two different S-R assignments). Unlike separate-coding approaches that invoke independent modules for stimulus and response processing, TEC leads us to expect that concurrent processing of the two events results in direct interactions, with the degree of interaction being modulated by inter-event similarity. Therefore, the major empirical questions are, first, whether performing a particular action is affected by the concurrently presented stimulus ($S_n \rightarrow R_n$) at all and, second, whether perceiving a particular stimulus is affected by the concurrently performed action ($R_n \rightarrow S_n$, as measured by performance of $R_{n+1}$). In fact, both kinds of effect were observed by Schubö and her colleagues.

First, it was found that the amplitudes and velocities of copying movements were clearly affected by the motion pattern monitored at the same time ($S_n \rightarrow R_n$), even though these two events did in no way depend on each other (i.e., $S_n$ was to be copied by $R_{n+1}$ and Rn referred to $S_{n-1}$). In particular, the amplitude and velocity of $R_n$ decreased with increasing amplitude and velocity of $S_n$, and vice versa. That is, $S_n$ had a contrasting effect on $R_n$. And this is precisely what TEC would have to predict: Perceiving the event $S_n$ requires the binding of its features, that is, the integration of the codes activated by presenting $S_n$. If $S_n$ is a fast motion, say, the feature code FAST becomes part of $S_n$'s object file and, thus, is no longer freely available for other bindings. However, if $R_n$ is also intended to be fast, this poses a problem for planning $R_n$, because the required FAST code is already occupied. As a consequence, a "less fast" parameter will be chosen, so that $R_n$ will tend to be slower as required.

Second, there was evidence that the memory of the event observed ($S_n$, as measured through $R_{n+1}$) was modulated by the event being generated ($R_n$) at the same time. This effect, however, was not an effect of contrast, but one of as-



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

similation or compatibility: The memorized (and later produced) amplitude and velocity tended to get smaller/slower when $S_n$ was accompanied by a small/slow movement and to get larger/faster when $S_n$ was accompanied by a large/fast movement. Again, such an outcome complies with TEC: Conflict and competition for feature codes between stimulus representations and action plans should only be observed as long as the competing stimulus- and action-related bindings are intact. However, from the study by Stoet and Hommel (1999; see above) we know that action plans desintegrate immediately after use, this leaving only some activation, but no longer occupation of the involved feature codes. Now, assume that in the task of Schubö and colleagues $S_n$ was of medium velocity while $R_n$ was rather fast. After $R_n$ was carried out, the corresponding action plan would desintegrate and leave some activation in the feature code FAST. This leftover activation would coexist for some time with the memorized MEDIUM code representing the velocity of $S_n$, which clearly invites an assimilation of the two codes. As a consequence, the stored codes belonging to $S_n$ would tend to be distorted in the direction of the codes used for performing $R_n$, which then shows up in assimilation effects on $R_{n+1}$ just as found.

Similar aftereffects of action on perception or, more precisely, of manual response selection on the selection of visual stimuli, can be taken from the experiments by Hommel and Schneider (in press). In their study, participants performed two temporally overlapping tasks, a speeded manual tone discrimination and an unspeeded visual bar-probe task. One of the interesting variables was the spatial relationship between $R_1$, the manual response, which was a left or right keypress performed with the index or middle finger of the right hand, and $S_2$, the marked letter target, which could appear on the left or right side of a four-letter display. As it turned out, spatial $R_1$-$S_2$ correspondence had a reliable and consistent effect on performance in the bar-probe task: Reports of left-side targets were more often correct following a left rather than a right manual response, and reports of right-side targets were better following a right than a left response. This observation allows for two interesting conclusions. First, response-selection can affect stimulus-selection, which strongly suggests that processes operate on the same cognitive codes within the same representational medium. Second, this impact of response on stimulus selection is positive, not negative, that is, it is associated with benefits of feature overlap, not costs. This fits well with the assumption that the aftereffects of feature activation and feature binding behave very differently in perception and action planning: While in perception binding effects outlast the effects of activation, the opposite is true for action planning.

### 4.4. Task-related modulation of event coding

One of the most interesting common characteristics of perception and action planning is that both are selective: We do not attend to all perceivable features of a given object and do not intend each of the many outcomes a particular action produces – at least not to the same degree. In TEC this kind of selectivity is modeled through the differential weighting of stimulus- and response-feature codes. Thus, we assume that task-relevant features of perceptual or action events are weighted more strongly than irrelevant features, so that they are more likely to be integrated into event codes and, thus, define the to-be-coded event, without necessarily excluding all irrelevant information.

Some evidence for selective coding and integration can be taken from the studies on object files. The original object-file approach of Kahneman et al. (1992) assumes that all features of an attended object are equally represented in an object file, except its location, which serves to address and access the corresponding file. However, we have already mentioned that object-specific preview benefits survive changes of letter case (Gordon & Irwin 1996; Henderson 1994) or horizontal orientation of pictures (Gordon & Irwin 1998), which speaks against Kahneman et al.'s (1992) complete-integration assumption. Furthermore, Hommel (1998b) observed preview benefits only for conjunctions of form (a task-relevant stimulus feature) and color (a task-irrelevant stimulus feature), and of form and location, but neither the conjunction of color and location nor the triple-conjunction of form, color, and location produced an effect. In other words, there was evidence for the binding of non-spatial stimulus features – independent of task relevance – and for the binding of location and the task-relevant stimulus feature, but no further (i.e., more complete) bindings or dependencies were obtained. Interestingly, the same task-irrelevant features that showed less evidence for binding were also less likely to produce simple feature-repetition effects. Inasmuch as the latter can be attributed to mere code activation (in contrast to integration), this suggests that it is the degree of code activation that determines the likelihood to be involved in feature-integration processes, just as TEC proposes. Thus, it seems that both the stimulus-induced activation and the integration of features in visual perception are incomplete, but are modified by task-specific feature selection or weighting.

Very similar observations have been made in experiments on stimulus response binding. As already indicated, Hommel (1998b) found strong evidence for the task-specific coding and integration of stimulus and response features. If form was task-relevant, repetitions of form and response or alternations of both produced better performance than incomplete repetitions, suggesting that response information was integrated with stimulus form. At the same time, color and its relationship with the response did not play a role, which points to a task-specific feature-integration process. In fact, simply redefining the task by making color relevant and form irrelevant completely reversed the result pattern: Now the response was integrated with color, but not (or not that much) with form.

Feature weighting is not only observed in the integration of stimulus features or of stimulus-response combinations, but in action coding as well. Recall that in the study by Hommel (1993), participants made pitch-discrimination responses by means of pressing a left-vs-right key, which switched on a light on the opposite side. One group of participants was instructed in terms of "key-pressing" and another in terms of "light-flashing." In the key group performance was better with stimulus-key correspondence, suggesting that the responses were coded in terms of key or effector location, while in the light group performance was facilitated by stimulus-light correspondence, thus, responses were coded in terms of the experimentally induced action effect. Note that the stimuli and responses were identical in both groups, yet the response feature (i.e., its perceivable effect) that dominated the action's cognitive representation clearly depended on the instructed action goal, that is, on


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

the subjective task relevance of the action's outcomes. Obviously, the same action feature can play a very different role, the importance of which is a direct function of the cognitive representation of the task at hand.

Still, task goals only modulate, but do not exclusively define the representations of task-specific actions. In fact, there is evidence that, although codes of intended action effects are rather dominant, nonintended, irrelevant effects are also integrated into action files. For instance, Hommel (1993: Experiment 2) did not only observe the strong effect of stimulus-goal correspondence just described, but also (much smaller) contributions of the correspondence between stimulus and (anatomically defined) hand and between stimulus and hand position, both independent of the goal-related effect. This indicates that action features that are unrelated to the defined, intended action goal may be weighted less strongly than the intended ones and, thus, get integrated more likely and/or more easily into the action's event code, but they can also become coded and integrated to some degree. This is also shown by the findings from another study (Hommel 1996b: Experiment 1), where people pressed left and right keys in response to color stimuli that randomly appeared on the left or right side. In one part of the experiment, each keypress produced a tone on the opposite side, similar to the design of Hommel's (1993) intention study, so that stimulus-key correspondence was associated with stimulus-tone noncorrespondence and vice versa. Although participants were instructed in terms of response keys, the presence of the auditory action effect markedly reduced the benefit of stimulus-key correspondence. That is, the noncorrespondence between the stimulus and a task-irrelevant, but hard-to-ignore action effect must have worked against the facilitative effect of stimulus-key correspondence (and vice versa). This provides strong evidence that the irrelevant, learned effect was integrated into the actions' representations, but to a lesser degree than the relevant effects.

At first sight, it may appear that the automatic integration of irrelevant action effects is not an overly reasonable and parsimonious way to represent an action, because it only seems to overload the cognitive system with unnecessary information. However, at second sight the automatic integration of currently irrelevant action effects turns out to be the very basis of intentional performance (Hommel 1997; 1998a). Actions are intentional activities by definition and therefore require an anticipated action goal, that is, an intended action effect. However, performing an action in order to produce an intended effect presupposes advance knowledge about the respective action-effect relationship. This again indicates that the acquisition and integration of an action-contingent event must precede the actual use of its code in planning and performing the associated action or, in the words of James (1890, p. 487), "if, in voluntary action properly so-called, the act must be foreseen, it follows that no creature not endowed with divinatory power can perform an act voluntarily for the first time."

Intentionally acting thus implies that the action is selected by (activating codes of) its consequences, so that the original learning experience "action → effect" is in some sense reversed to "effect-code activation → action." In a recent study, Elsner and Hommel (2001) attempted to mimic this reversal in an experimental situation. In an acquisition phase, participants were presented with a simple trigger stimulus on which they were required to press one of two keys, whichever they wished (i.e., free choice). Pressing one of the keys produced a high tone and pressing the other a low tone, both being completely irrelevant to the task. Then there was a transfer phase consisting of a binary choice RT task with the high and low tone as relevant stimuli. One group received a "compatible" mapping, thus, the tones were assigned to the responses they had followed in the acquisition phase, while the other group had an "incompatible" mapping, with tones assigned to the responses they had not followed. From a theoretical point of view, this situation should be very similar to the one perceivers/actors are faced with when learning to act in an intentional way. First they need to acquire the contingency between a particular movement and its perceivable outcome (i.e., form a movement-effect association), and then to use the code of the outcome to intentionally select the associated action. If so, the participants in the Elsner and Hommel study should have acquired associations between responses and tones, which then should prove to be useful in selecting the correct responses during the transfer phase in the compatible, but not the incompatible, group. The effect of compatibility on performance was dramatic, indeed, with the compatible group by far outperforming the incompatible group. Interestingly, large compatibility benefits were observed even if the tones were no longer presented as action effects during transfer, indicating that tone information was still maintained in the actions' representations.

Related evidence for an interplay between intentional and stimulus-triggered processes comes from studies on sympathetic action, commonly called ideomotor phenomena. Only recently, Knuf and colleagues (Knuf 1997; Knuf et al. 2000) developed a task that allows for an experimental analysis of ideomotor action. This task was modeled after the classical bowling example (Lotze 1852): Consider an observer who has just set a bowling ball going and is now following its course. As long as it is open whether it will reach its intended location or not, he or she can hardly prevent him/herself from moving his/her hand or twisting his/her body as if to exert some kind of magical impact on the ball. In the experimental task, on each trial a ball, starting from the bottom of a computer screen, would move toward a target, which was visible at a certain position at the top of the display. During an initial instrumental phase, participants were able to influence either the path of the ball or the horizontal position of the target by means of a joystick, thereby getting the chance of hitting the target which would never have happened without correction. More important was the subsequent induction phase, where, as the participants knew, joystick movements no longer had an effect. Nevertheless, the joystick was monitored for (now dysfunctional and useless) movements, which were then analyzed as a function of the ball's path, of whether the target was hit or missed and, if applicable, on which side the target was missed.

From a TEC perspective, one may come up with two sets of predictions, one that does not consider a special role of intended action effects and another that does. According to the *perceptual induction view,* perceived events will induce similar actions (cf. James 1890, Vol. II, p. 526), suggesting that joystick movements during the induction phase would always follow the motion of the ball. According to the *intentional induction view,* however, intended events may be more critical than perceived events, so that joystick movements during the induction phase should critically depend



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

on whether the ball is expected to hit or to miss the target. If a hit is considered, no ideomotor movements are expected at all. However, if a miss is foreseen, this should induce movements that would be suited to achieve a hit. In the ball condition (where the ball is under the observer's control), this would be attempts to push the ball toward the target; in the target condition (where the target position is under control), it would be attempts to push the target toward the ball.

Taken together, the results of Knuf and colleagues clearly rule out a pure perceptual induction account, but lend substantial support to the intentional induction view. First, there was no substantial indication that the direction of ball motion was a major determinant of the direction of induced movements. Second, induced movements were virtually absent on trials resulting in target hits. Third, a great number of induced movements were observed on trials resulting in misses, with a direction that depended on experimental condition (ball vs. target) and side of target miss, in the way predicted by intentional induction. In conclusion, ideomotor movements are more strongly governed by representations of intended rather than perceived events, at least in Knuf's paradigm. Therefore, ideomotor phenomena should be considered derivatives of instrumental, goal-directed action rather than derivatives of spontaneous imitation. Again, this does not completely rule out contributions from nonintended events. In fact, further experiments by Knuf and colleagues have revealed that perceptual induction may be quite effective under certain conditions. For instance, while effectors that are instrumental in a given task (i.e., hands) exhibit signs of intentional induction only, an analysis of ideomotor (dysfunctional) movements in non-instrumental effectors, such as head or feet, showed evidence of perceptual induction superimposed on intentional induction. Such a pattern fits well into the general picture: Intended action effects and their codes play a central role in action control, but this does not exclude some limited contribution from nonintended, currently irrelevant effects – presumably because it is always possible that these represent the actor's future goals.

### 4.5. Summary

Taken together, we have seen that the basic principles and assumptions underlying TEC are not only logically tenable, but also well supported by the empirical evidence. First, there are numerous examples of interactions between perception and action planning, going in one or the other direction, and it is hard to see how these could be sufficiently well explained without assuming that perceptual contents and action goals are coded within the same, common representational medium. Second, we have seen that what matters for the interplay between perception and action planning is the similarity between perceived and to-be-produced events, that is, the feature overlap between perceptual events and action plans. Obviously, this suggests that perceived and produced events are cognitively represented not as unitary wholes, but as bundles of feature codes, just as TEC presumes. Third, and in accordance with TEC, there is strong evidence that the feature codes that represent a given perception or action event do not refer to proximal, sensory or muscular regularities, but to distal attributes of the event, such as the perceived location of an external movement-contingent feedback light.

There is also good support for TEC's more specific assumptions about the anatomy and characteristics of event codes. First, it seems clear that some codes or parts of codes are involved in both perception and action planning, so that mutual access to, or use of, these codes brings about characteristic patterns of facilitation and interference between perception and action. Second, it is also evident that those patterns are nicely predicted by making a distinction between the activation and the integration of codes, and by assuming that these two processes have very different effects on concurrent perceptual or action-planning processes. And finally, there is increasing support for the assumption that the activation and integration of feature codes is modulated by task relevance and task-related intentions.

We have presented converging evidence from a wide variety of tasks, including single keypresses with hand, foot, and tools; keypressing sequences; verbal actions; rotatory, pointing-type, and positioning movements; hand gestures; drawing movements; and sensorimotor synchronization. Of course, the investigated perceptual events and actions were often rather simple and arbitrary, reflecting their experimental origin. True, speaking, gesturing, and (increasingly) keypressing fills considerable parts of our daily life, so that we are optimistic to already have captured areas of relevant and representative human behavior. Yet, it is also clear that perceptual events can be more complex than single flashes on a screen and action plans can refer to something more complicated than finger movements. In principle, this does not change the picture, because most complex perceptual and action events are merely combinations or sequences of the components investigated in the experiments discussed above (see sect. 3.2.5). That is, there is no a priori reason to assume that TEC might not be applicable to arbitrarily complex perceptual events and action plans.

The central strength of TEC is, in our view, not so much that it can easily account for the phenomena and effects discussed here; after all, one could well imagine that one or the other competing approach could be extended or modified to account for at least some of these findings. More importantly, TEC has inspired us to study a number of phenomena that would not have been expected and not even been looked for from other perspectives. Thus, apart from its explanatory success, it is the heuristic value of TEC we wish to stress, its potency in stimulating new questions and research strategies – even if answering these questions and following these strategies will very likely result, somewhere in the future, in the replacement of TEC by some other, more elaborate and detailed theory of the relationship between perception and action planning.

## 5. Epilogue

As a concluding remark we would like to discuss an argument of one of the classical issues in epistemology: How is it possible that our mental experience refers to events in the outer world (in the distal environment), given the fact that it derives from events in our bodies and our minds (i.e., from proximal stimulation and central processing)? In other words: Why does the mind construct objective, external reality and how does this happen?

The argument we offer is based on the functional logic inherent in TEC. It takes two steps. First, we suppose that, at some (unknown) point in the evolution of animal minds,


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

it was highly advantageous to develop a new representational domain for matching perception and action planning – on top of the old system for mapping and response selection. Second, we hold that the most efficient way of implementing a domain like this is by creating codes with distal reference, that is, codes whose content refers to events and actions in the animal's environment. In a nutshell, then, the argument is that the invention of identical codes for perceived and to-be-produced events implies the construction of reality.

### 5.1. Inventing identical codes

We presume that common codes have evolved as a means to find a transition from responding to given stimulus conditions to planning goal-directed actions. Consider, as a thought experiment, an animal that wishes to complement its old and simple device for response selection by another, more sophisticated device for action planning. What kinds of functional investments need to be undertaken to solve this adaptive problem? Basically the planning device needs to bridge two functional gaps, between events and actions and between present and future.

First, as the animal's actions need to be planned in accordance with the course of events in its environment, the tuning of future action to the forthcoming pattern of environmental events requires a representational scheme for commensurate coding of events and actions – not only in terms of spatial and temporal coordinates but of semantic content as well. A scheme like this is not required for response selection. This is because response selection is an on-line operation, selecting reactions in response to actually given stimuli. In contrast, action planning is an off-line operation, planning future actions in accordance with expected and/or intended future events.

This brings us to the second functional gap. The action planning device needs to find a way to bridge the gap between events going on at present and events forthcoming in the future, that is, between on-line generated perceptual representations and off-line generated conceptual constructions. Moreover, representations of absent future events can play two major functional roles: expectations and intentions. For goal-directed action planning, intended events, or goals, need to be tuned to expected events.

The common representational domain invoked by TEC offers a solution for bridging both gaps. When viewed in terms of its adaptive achievements it is ideally suited to serve the needs of an action-planning device. The principle behind the solution TEC offers is simple enough. Identical codes of perceived and to-be-produced events are created by way of transforming information from low (motor and sensory) to high (perceptual and conceptual) levels of representation. Events and actions that are incommensurate at the level of sensory and motor features will be commensurate at the perceptual level. Likewise, perceived present events and expected, or intended future events will be commensurate at the conceptual level.

### 5.2. Inventing reality

Why do we see a flower when we look at it? We could just as well see, for example, light rays, retinal excitation patterns, spikes traveling along our nerves or patterns of chemical or electrical activity in various parts of our visual cortex. Why don't we see all these closer links in the chain that leads from the flower to the brain? This question has been raised by a number of authorities, including Brunswik (1944), Dretske (1981), Gibson (1966), Heider (1926/1959; 1930/1959), Helmholtz (1909), and Köhler (1929). In Dretske's words the miracle is this: Our percepts carry "information about a distant causal antecedent . . . without carrying information about the more proximal members of the causal chain . . . through which this information . . . is communicated." It is as if the content of our perceptual experience "skips over (or 'sees through') the intermediate links in the causal chain in order to represent (carry information about) its more distant causal antecedents." (Dretske 1981, p. 158). Dretske argues convincingly that we must distinguish between a given percept's causal antecedents (of which there is a long chain) and its informational content (which refers to one particular link in this chain). Yet, this distinction does not speak to the question why in this complex network of antecedent conditions the particular layer of external objects and events is singled out. Why should we see the flower – and not the spikes in our brain?

The principle of common coding inherent in TEC offers a straightforward answer to this query (cf. Prinz 1992): The reason why we perceive distal objects and events (and plan our actions in terms of distal events as well) may be that it is only at the level of distal representation that event codes and action codes can be generated and maintained in a common format. Likewise, the same may apply to the present and the future: It may be only at the level of distal representation that perceptual representation and conceptual construction can meet. Therefore, the format of distal events may be singled out from the causal chain because it is the sole format that permits commensurate coding and, thus, allows for efficient action planning.

Why should this be the case? There are two reasons to believe that action planning is dependent on distal representations and does not work with proximal representation. First, coordinating actions and events is much easier in terms of distal than proximal representations. Proximal representations are both incommensurate and incomplete, one class of representations referring to sensory features and the other to motor features – and none of them referring to semantic features at all. Any action planning on this basis would therefore require heavy interaction between these three domains of representation. Conversely, with distal representations, the coordination of actions and environmental events can be performed in a single and coherent domain.

Second, and more important, the forecasting of events is much easier in terms of distal than proximal representations. Coordinating future actions with future environmental events requires representational resources capable of (1) elaborating reliable scenarios for the future course of environmental events as well as (2) building one's own actions into this course in order to achieve certain goals. For obvious reasons these two tasks are hard, if not impossible to solve with proximal representations. Forecasting future and environmental events in terms of their proximal (sensory) representations would imply an explosion of degrees of freedom, due to the virtually infinite number of the observer's possible perspective on those events. The same would apply to the planning of future actions in terms of their proximal (motor) representations: There is always a



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

virtually infinite number of bodily movement patterns through which a given action (specified as a meaningful environmental event) can be realized. As a consequence there is no reasonable way to plan the future in terms of proximal representations. Forecasting events and planning actions requires representational resources with distal reference. It is important to note that, in this view, it is not the case that common coding follows from distal representation. Instead we believe that abstract, distal representation has evolved as a solution to the problem of developing a representational scheme for the planning of goal-directed action. Action planning has been the problem, common coding has been the solution, and reality has come as a by-product of the solution.

ACKNOWLEDGMENTS

We wish to thank Bruce Bridgeman, Diego Fernandez-Duque, Peter H. Greene, Anthony Greenwald, Dan Lloyd, Claire Michaels, Robert E. Shaw, and the anonymous reviewers for valuable comments, criticisms, and suggestions on earlier versions. Thanks are also due to Heidi John for stylistic suggestions. Requests for reprints can be sent to Bernhard Hommel, University of Leiden, Section of Experimental and Theoretical Psychology, Wassenaarseweg 52, P.O. Box 9555, 2300 RB Leiden, The Netherlands or Jochen Müsseler, Gisa Aschersleben, or Wolfgang Prinz at the Max-Planck-Institut für Psychologische Forschung, Amalienstraße 33, D-80799 München, Germany. Email: hommel@fsw.leidenuniv.nl, muesseler@mpipf-muenchen.mpg.de, aschersleben@mpipf-muenchen.mpg.de, prinz@mpipf-muenchen.mpg.de.

NOTES

**1.** This requirement is closely related to the notions of inverse kinematics and inverse dynamics in motor control (Rosenbaum 1991). In its most abstract form the underlying problem can be stated as follows: Given a certain intended event, or goal, how can body movements be organized to effectuate that event? In the motor literature this problem is applied to the physics of action, that is, dynamics and kinematics. In ideomotor theory the issue is generalized to also include the semantics of action.

**2.** This hypothesis implies a much stronger role for goal representations than is usually implied in theories about the role of feedback in motor learning and motor control. These theories, too, hold that representations of intended action effects are maintained ("Plans"; Miller et al. 1960; "Image of Achievement," Pribram 1971). However, their main function is typically seen to be related to evaluation (i.e., testing observed against expected outcome), not to action preparation and control (Hommel 1998a).

**3.** The distal-coding principle distinguishes our approach from those of Greenwald (1970), James (1890), and Lotze (1852), who all focused on the proximal feedback from movements. In our view, it is this rather tight focus that stood in the way of a broader impact of these approaches. The domain of stimulus-response compatibility provides a good example (for overviews, see Hommel & Prinz 1997; Proctor & Reeve 1990). As Greenwald has argued, Lotze and James's ideomotor principle represents a promising starting point for a theoretical account of most compatibility phenomena observed so far. However, researchers have (more or less correctly) interpreted both the original principle and Greenwald's learning-theoretical version as being restricted to the very rare cases of proximal similarities between stimuli and responses (e.g., Gordon 1990; ten Hoopen et al. 1982). Given that most compatibility phenomena are based on more abstract and sometimes purely symbolic stimulus-response relationships, such as between direction words and spatially directed movements, this means that an ideomotor approach is often not applicable. However, once the ideomotor principle is combined with the assumption of distal coding, it changes into a potentially powerful theoretical tool (Hommel 1997; Prinz 1990).

**4.** Actually, we do not believe that uttering the word "left" and moving one's hand leftward are necessarily controlled by exactly the same code. More plausible would be that both action features are first represented by different and independent codes. However, through learning to use the word "left" to refer to things that can be pointed at with a leftward hand movement (i.e., learning to co-activate word- and hand-related codes) these two codes get tightly associated, so that activating one of them will spread activation to the other. To the degree that this spreading of activation is automatic, the activated pair of codes can be regarded as one functional code. At any rate, our framework does not hinge on whether the reader will follow us at this point.

**5.** There are subtle differences between the models as to exactly how differences in processing time are produced: According to the so-called "Paillard-Fraisse hypothesis" differences in nerve conduction times between click and tap on their way to their central representation are responsible for the anticipation error (Aschersleben & Prinz 1995; 1997), while Gehrke's "sensory accumulator model" (Gehrke 1996; Aschersleben et al. 2000a) assumes that clicks and taps differ with respect to the amount of sensory evidence required for central coding.

**6.** Of course, depending on the task demands, a supplementary impairment of actions by concurrent stimulus processings should also be obtained. In fact, such an impairment was observed in a task where participants performed a sequence of left-right keypresses during the presentation of right or left-pointing arrows. Critical keypresses were delayed when executed during the presentation of an action-compatible stimulus as compared to an action-incompatible stimulus (Müsseler 1995; see also Prinz 1997a).

# Open Peer Commentary

*Commentary submitted by the qualified professional readership of this journal will be considered for publication in a later issue as Continuing Commentary on this article. Integrative overviews and syntheses are especially encouraged.*

## Intelligent control requires more structure than the Theory of Event Coding provides

Joanna Bryson

*Computers and Cognition Group, Franklin W. Olin College of Engineering, Needham, MA 02492.* **joanna@ai.mit.edu** **http://www.ai.mit.edu/~joanna**

**Abstract:** That perception and action share abstract representations is a key insight into the organization of intelligence. However, organizing behavior requires additional representations and processes which are not "early" sensing or "late" motion: structures for sequencing actions and arbitrating between behavior subsystems. These systems are described as a supplement to the Theory of Event Coding (TEC).

**Hommel et al.** have made a substantial contribution to our understanding of the interrelation of perception and action. Considering that well-ordered action is the selective advantage of intelligence, it is clear that perception is fundamentally *a part* of action. Given our understanding of neural learning, it is also not surprising that as regularities in coding become established, either in perception or in motor control, they will become linked directly to any other well-correlated systems that happen to be able to monitor their activity. It is generally held that this explains the separation of the sensory and motor cortical maps in all but the most primitive mammals: separation provides greater opportunity for



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

the intervention of between-process action control (Livesey 1986).

There is, however, more to action representation and control than what the Theory of Event Coding (TEC) currently provides. Also, this supplementary control is often not explicit or intentional, despite being at a higher level than the TEC representations the authors claim are necessarily intentionally accessed. Perception and action are also sometimes fused even more tightly than feature-based coding implies. In this commentary, I address two forms of action control beyond those specified in TEC: the sequencing of basic actions, and arbitration between possible higher-level behaviors or goals.

That sequences of actions are themselves a privileged representation has been well supported for some time in the neuroscience literature (Houghton & Hartley 1995; Lashley 1951). Many expressed action sequences are too quick and intricate for perceptual feedback to control the transitions between them. Further evidence of privileged representations for behavior patterns is expressed by experiments where animals with their forebrains removed are still capable of conducting complex species-typical behaviors, although they are unable to apply these behaviors in appropriate contexts (Carlson 2000). In particular, the periaqueductal grey matter has been implicated in complex species-typical behaviors such as mating rituals and predatory, defensive, and maternal maneuvers (Lonstein & Stern 1997).

Evidence of brain cells directly involved in processing sequential recall and behavior has also been found by Tanji et al. (1994). Tanji and Shima studied trained monkeys performing sequences of actions either from memory or by following cues. They recorded cells firing in the medial frontal cortices in three different functional contexts: (1) during a particular action, regardless of whether it was recalled or cued; (2) between particular action pairs, only when recalled, but regardless of which sequence they occurred in; and (3) between signal and initial action for a particular complete sequence when it must be performed from recall. These cell types made up 12%, 36%, and 26% of the cells recorded, respectively. The latter two types indicate special coding for types of transitions and full sequences of behavior.

Graziano et al. have also recently shown that the premotor and motor cortices contain single cells representing not only complex motion sequences, but also multimodal triggering of these sequences. Example behaviors generated by these cells include feeding (making a grasp at a particular location in egocentric space and bringing the hand to an opening mouth), scratching and ducking as if from blows or projectiles (simultaneously raising an arm and averting the head) (Graziano & Gandhi 2000; Graziano et al. 2001). The multimodal perceptual input is often keyed to a particular location relative to the relevant organ, for example, the head for ducking, or a limb to be scratched, and can be triggered by sights, sounds, or tactile sensations. The action-sequencing cells seem to be organized in the same roughly topographic mapping (with respect to the location of the triggering stimuli) common in perceptual cortices from V1 through face orientation cells in the temporal cortex (Perrett et al. 1992).

Even if Graziano's cells reference common feature mapping for sensing and action in the way that TEC implies (which I doubt), it seems unlikely that any intentionality is intervening between stages of this process. By this I do not mean to say that TEC-like representations do not exist, rather that perception and action are probably unified in a number of different ways, of which TEC is but one.

Another is the level of arbitration between behaviors: the problem mentioned earlier of choosing appropriate contexts in which to act, or behavior processes to attend to. Arbitration is necessary in a parallel, distributed model of intelligence because an agent has finite resources (e.g., hands and eyes) which the different possible behaviors (e.g., feeding and mating) must share. Arbitration must take into account both the activation level of the various "input" cortical channels and previous experience in the current or related action-selection contexts.

Recently, the basal ganglia (BG) has been proposed as the system responsible for this aspect of action coordination (Gurney et al. 1998; Mink 1996; Prescott et al., forthcoming; Redgrave et al. 1999). The BG is a group of functionally related structures in the forebrain, diencephalon, and midbrain. Its main "output" centers – parts of the substantia nigra, ventral tegmental area, and pallidum – send inhibitory signals to neural centers throughout the brain that either directly or indirectly control movement, as well as other cognitive and sensory systems (Middleton & Strick 2000), which is fitting for an arbitration system. Its "input" comes through the striatum from subsystems in both the brainstem and the forebrain, giving it access to both automatic and "intentional" cues for redirecting attention.

Finally, although my commentary has focused on evidence from neuroscience, my understanding of the roles of sequence, arbitration, and sense/action coupling derives from my own experience with artificial intelligence. For 15 years, behavior-based AI (BBAI) has built intelligence by explicitly combining perception and action into uniform representations, producing systems capable of complex, animal-like behavior. Although the original BBAI systems were fully distributed (Brooks 1986), difficulty in scaling the complexity of such systems have led the field toward incorporating control structures such as sequence and arbitration for coordinating the actions of the behaviors (Bryson & Stein 2001; Kortenkamp et al. 1998). I encourage readers to explore this literature for a further understanding of intelligent control.

ACKNOWLEDGMENTS

Thanks to Will Lowe for his comments on an earlier draft. Parts of this document were generated using the LaTeX2HTML translator Version 99.2beta8 (1.42).

## A common framework for perception and action: Neuroimaging evidence

Thierry Chaminade[a] and Jean Decety[b]

[a]*Neurophysiology of Intentionality, INSERM Unit 280, Lyon, 69 424 Cedex 3 France;* [b]*Center for Mind, Brain and Learning, University of Washington, Seattle, WA 98195.* **chaminade@lyon151.inserm.fr**
**http://www.lyon151.inserm.fr/unites/280_1angl.html**
**decety@lyon151.inserm.fr**
**http://www.lyon151.inserm.fr/unites/280_1angl.html**

**Abstract:** In recent years, neurophysiological evidence has accumulated in favor of a common coding between perception and execution of action. We review findings from recent neuroimaging experiments in the action domain with three complementary perspectives: perception of action, covert action triggered by perception, and reproduction of perceived action (imitation). All studies point to the parietal cortex as a key region for body movement representation, both observed and performed.

***Introduction.*** The Theory of Event Coding (TEC) described in **Hommel et al.**'s target article is based mainly on psychological data. We believe that the neurophysiological level deserves to be addressed as well in this theory. The lack of discussion of neurophysiological data in the TEC framework is surprising, given that in the last decade strong evidence from neuroimaging studies has accumulated in favor of a common representational domain for perceived and intended, or to-be-generated, events. This has led to the proposal that neural simulation of action, comprising intending, preparing, imagining, and observing actions, should be considered as a unifying mechanism for motor cognition (Jeannerod 2001). We believe, from our understanding of Hommel et al.'s target article, that these shared representations of actions belong to the "event codes" they describe.

We will briefly review some of the recent neuroimaging studies focusing on three sets of data belonging to the action domain, namely, action perception, action triggered by perception, and reproduction of perceived action (imitation), which have been addressed in the target paper. We think that the exploration of the



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

neurophysiological level is of major importance for validating the TEC as it can demonstrate the commonality of the brain structures involved in perception and generation of actions, the neural aspect of the shared representations of action. Moreover, this exploration can reveal the functional macroanatomy of these action-related "event-codes" which may provide knowledge of some aspects of their content.

***Perception of others' actions.*** One impressive illustration of the functional role of shared representations between perception and action is provided by the so-called "apparent motion phenomena," that is, the subjective visual impression resulting from the sequential presentation of static objects in different spatial locations. When presented with sequential static images of an inanimate object in different positions, the object is perceived as moving along the shortest or more direct path even when such a route would be physically impossible, for example, when it would require one object to pass through another object (Shiffrar 2001). Thus, the visual system seems to be biased toward selecting the simplest interpretation of the image when it involves inanimate objects. However, the perception of apparent motion operates differently when the object presented is a human figure, as demonstrated by the extensive work by Shiffrar and Freyd (1990; 1993). These authors have shown that viewing images depicting apparent motion of humans, depending on the time interval between the stimuli, gives rise to the perception of either a direct path (i.e., biomechanically implausible) or an indirect path (biomechanically plausible). This paradigm was used by Stevens et al. (2000) to investigate, with Positron Emission Tomography (PET), whether a difference in brain activity accompanies this perceptual difference. They found neural encoding of apparent motion to be a function of the intrinsic properties of the stimulus presented (object vs. human), as well as the kind of human movement path perceived (plausible vs. implausible). Premotor and inferior parietal cortex were only involved in the perception of biomechanically possible motion, suggesting that these regions are selectively activated to process actions that conform to the capabilities of the observer.

Numerous neurophysiological experiments have shown that observing actions performed by others activates parietal and premotor regions (for a meta-analysis, see Grezes & Decety 2001). The perception of body movements elicits hemodynamic responses in a somatotopic manner – simply watching mouth, hand, and foot movements activates the same functionally specific regions of premotor cortex as making those respective movements (Buccino et al. 2001). It has indeed been recently demonstrated that this motor resonance phenomena occurs only for real human actions. Perani et al. (2001) investigated whether observation of actions reproduced in three-dimensional virtual reality would engage perceptual and visuomotor brain processes which are different from those induced by the observation of real hand actions. Subjects were asked to observe the grasping of geometrical objects performed either by a real hand or by hand reconstructions of different quality in 3D virtual reality as well as on a 2D TV screen. The authors found that only real actions in natural environments activated a visuospatial network including the right posterior parietal cortex and the left parieto-temporal region. Observation of virtual-reality hand actions engaged prevalent visual perceptual processes within lateral and medial occipital regions. Thus, only perception of actions in actual reality maps onto existing representations, whereas virtual-reality conditions do not access the full motor knowledge available to the central nervous system.

The work of Kandel et al. (e.g., 2000) is cited in the target article to exemplify the links between perception and action. Their work shows that perception of an action, in this case the writing of a letter, allows one to anticipate the forthcoming letter. A similar "perceptive anticipation" paradigm was recently adapted into a PET study by Chaminade et al. (in press a). Stimuli consisted of a single moving dot whose trajectory depicted the beginning of either mechanical, pointing, or writing movements. The mechanical movement represented a ball bouncing on a spring and going either far or near, whereas pointing corresponded to a finger movement toward a big or small target, and writing to the handwriting of the first *1* in a pair of letters, either *ll* or *ln*. A control condition presented the dot either at the top or at the bottom of the screen. PET measurements of regional cerebral blood flow were conducted in eight healthy subjects who were requested, in a forced-choice paradigm, to anticipate the next step of the presented event. Our hypothesis was that brain systems engaged in anticipating the continuation of an observed action would overlap brain systems used for the execution of the same action. Results showed that anticipation of pointing specifically activated the right intraparietal sulcus and left premotor area, both involved in preparing a pointing movement, and anticipation of writing similarly activated Broca's area and the left superior parietal lobule, both involved in writing. Different motor subcircuits were therefore specifically recruited in the two tasks involving anticipation of observed human action.

Results discussed in this section demonstrate that perception of action relies on brain structures specifically implied in the production of the same action. They corroborate the idea that perception of human movement is constrained by the observer's knowledge of, or experience with his or her own movement limitations (Shiffrar & Freyd 1993), as well as Hommel et al.'s theory about the existence of shared representations of action, or "event-codes."

***Covert action triggered by perception.*** Behavioral studies have explored the existence of an automatic link between the visual perception of an object and a specific predetermined action, as illustrated by Gibson's affordance principle (sects. 2–3 of the target article). Tucker and Ellis (1998; Ellis & Tucker 2000) have performed an elegant series of experiments with a stimulus-response compatibility paradigm using photographs of common graspable objects as stimuli. The aim of their study was to examine the relation between an object's orientation and the hand most suited to perform a reach-and-grasp movement. Objects were presented in two horizontal orientations (compatible with either right-hand grasp or left-hand grasp) and two vertical orientations (upright or reverse). Subjects were requested to decide as fast as possible whether each object was upright or inverted. They reported faster right-hand responses when the object orientation was compatible with a right-hand grasp, and likewise for left-hand responses. These behavioral results are thus consistent with the view that seen objects automatically potentiate components of the actions they afford, even in the absence of explicit intentions to act. Grezes and Decety (in press) adapted this paradigm into a PET activation study and found left hemispheric neurovascular increases in the inferior parietal lobule, the SMA-proper, the inferior frontal gyrus, the dorsal and ventral precentral gyrus, as well as in the cerebellar hemisphere ipsilateral to the grasp side specific to the observation of objects. Such a pattern of activation is congruent with the idea that motor representation is already involved during the perception of an object, and thus provides evidence that the perception of objects automatically affords actions that can be made toward them.

Another way to induce covert action by means of perception is to instruct subjects to watch actions in order to imitate them later. PET was used to map brain regions associated with the observation of hand actions when subjects either had to recognize or imitate the seen actions (Decety et al. 1997). Observing with the intent to recognize activated the ventral stream of visual perception. In contrast, observation with the intent to imitate was associated with a strong increase of activity in the dorsal pathway extending to the premotor cortex bilaterally as well as to the cerebellum. Perception for later imitation therefore engages a network found in motor imagery and execution of action. Thus, perception of action already activates neural structures involved in motor planning of actions to be imitated later.

Both these examples demonstrate that neural simulation of action can be elicited by different percepts, for example, by the ob-



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

servation of someone else's action or of an object. This result is in line with Hommel et al.'s claim that "perceiving" an action or an object "and action planning are functionally equivalent." The next section shows how imitation paradigms can help to unravel the links between perception and action.

***Imitation of action.*** Imitation is a natural mechanism involving perception-action coupling which plays a central role in child development. It also provides a natural mechanism for studying the functional links between perception and action. So far, few neuroimaging studies have explored the neural correlates of different "levels" of imitation, from a "Stimulus-Response Compatibility"-like experiment (Iacoboni et al.1999), to the imitation of the higher-level goal of the action (Chaminade et al., in press b), but their results are promising.

Iacoboni et al. (1999) used functional Magnetic Resonance Imaging (fMRI) to test the "direct matching hypothesis" which postulates that imitation is based on a mechanism directly matching the observed action to an internal motor representation of that action. Volunteers were asked to perform a finger movement triggered either by the observation of a similar movement from a visible hand – thereby imitating the finger movement, – or by a symbolic cue. It was predicted that areas activated during finger movement should be more activated when a congruent movement is observed. Three regions were found to react that way, Broca's area, the left superior temporal sulcus (STS), and the right superior parietal lobule.

In our experiments, PET was used to allow imitation of more complex, object-oriented and goal-directed actions, which was not easily feasible in fMRI. The first study focused on reciprocal imitation (Decety et al., in press). Eighteen subjects either (a) imitated the experimenter's actions, (b) had their actions imitated by the experimenter, (c) freely produced actions, or (d) freely produced actions while watching different actions made by the experimenter. In a baseline condition, subjects simply watched the experimenter's actions. From a TEC perspective, it is important to consider that the same "visual" and "action" codes are present in conditions (a) and (b), so that they only differ by the higher-level event that ties them together. When subjects imitate the experimenter, "visual" codes must be translated into "action" codes, whereas in the reverse situation "visual" input is a consequence of action production. Specific increases were systematically detected in the left Superior Temporal Sulcus (STS) and in the inferior parietal cortex, as well as in the right frontal cortex in all conditions involving imitation compared to control conditions. These similar activation foci in a large number of brain areas show that the "visual" and the "action" codes are similar even though the "event," that is, the task, is different in the two conditions of imitation. At first glance, this could seem to go against the TEC which implies that all representations are modulated by the event, but one should remember that neuroimaging techniques only give a global view of the regional blood flow notwithstanding the computations performed in these areas. An alternative interpretation of our results is therefore that the similar brain areas found in the two contrasts – describing the two imitation situations – could perform different computations in the two conditions.

An interesting question raised in the target article is the "distality" of the event codes. Are they associated with low-level motor neural processing or higher-level goal representation? A second experiment was designed in order to study the relative roles of two important components of action, the end-point achieved by the action and the movements performed to achieve this action, during imitation of object-oriented actions (Chaminade et al., in press). Experimental conditions were derived from a $2 \times 3$ factorial design whose factors were: (a) the stimulus event shown to the subjects during the observation of the experimenter; and (b) the response manipulation performed by the subject. Two key components of human action, the goal and the means to achieve it, were systematically dissociated during the observation phase. Depending on the condition, subjects saw either a whole object movement performed by the experimenter, only its end-point or goal, or only the displacement of the object or means. During a response phase, subjects were asked either to copy the action they had just seen, either entirely or partially, or to act independently of the seen action. As expected from the previous experiment, imitation was associated with a bilateral inferior parietal and superior temporal activity. The important question concerned the differences in brain activity when imitation is triggered only by the observation of the means, or only by the observation of the goal. The results revealed partly overlapping clusters of increased activity in the right dorsolateral prefrontal area and in the cerebellum when subjects imitated either of the two components. Moreover, specific activity was detected in the medial prefrontal cortex during the imitation of the means, whereas imitating the goal was associated with activity in the left premotor cortex. The results were interpreted as follows: in both situations, imitation is based on a similar representation of the goal, which we postulate is located in the dorsolateral prefrontal cortex. However, when the end-point of the action is seen, the goal is directly available, but the actual motor programs need to be totally created, which accounts for premotor activity. On the other hand, when only the gestures associated with the action are shown, the goal must be found by reading the intention underlying the observed gesture, as implied by the medial prefrontal activity, whereas the motor program is already prepared by observation.

Our results therefore suggest that imitating a complex gesture activates processing of the intention (or goals) underlying the observed action, and cannot be described as a low-level resonance mechanism. Imitation is therefore a creative reconstruction of the observed action. This fits with the idea developed in the target paper (sect. 4.4) that ideomotor phenomena should be considered a derivation of goal-directed action instead of automatic imitation, or, in other words, that when imitating someone's action, the underlying intention or goal seems more important than the surface behavior itself (sect. 4.2.3).

***Concluding remarks.*** Inherent in the question of shared motor representations, and consequently, in the TEC as well, is the problem of action attribution. If "visual" and "motor" codes are not fully segregated at a certain level of task processing, how can we so easily differentiate between seen and performed actions? Quite surprisingly, this question is not addressed in the target paper. A first element can be found in the study of imitation described earlier. Contrasting the two conditions of reciprocal imitation shows the brain activity related to being, or not being, in control of the action performed. The left inferior parietal region is specifically involved in the imitation of the other by the self, whereas the right homologous region is more activated when the self imitates the other. A similar lateralisation of clusters of activity in this region was also found in another study in which subjects were asked to imagine either themselves or someone else performing given actions (Ruby & Decety 2001). Overall, these results favor the interpretation of a lateralisation in the inferior parietal region associated with self-related versus others-related information, in the dominant and nondominant hemispheres, respectively.

The results emphasize the multiple role of the parietal lobe, in particular its inferior part, in storing and using shared representations linking perception and execution of action. Parietal foci of activity were found in the vast majority of experiments in which action is either performed, simulated, or observed, and we have argued that it could be responsible for differentiating between action initiated by the self or the other. This region could play a key role in nonspecific shared motor representations. On the other hand, the study of perceptive anticipation shows that activation of these representations can also be specific to the task, for example, the writing system is activated when a decision has to be taken about a written item. This corroborates the TEC by showing that producing and perceiving a certain event share some specific activation foci. One last point we would like to make is about the cognitive level of the event-code. Results in the experiment studying imitation of the means or the goal of an action imply that there is a high-level, goal-directed representation of the to-be-performed



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

action. This also corroborates some aspects of the event-code hypothesis, where it is postulated that codes "represent th distal features of an event" (sect. 3.2.1).

Neuroimaging techniques are valuable tools for the understanding of behavior. Taking their results into account can only improve theoretical approaches based on psychophysical data, such as the one proposed in the target article, of mutual exchanges between theory and results. Though an extensive review of the neuroimaging studies related to the TEC was not the aim of this commentary, we believe that some insights from recent results focusing on the links between perception and action would be welcome in this discussion.

## Perception, action planning, and cognitive maps

Eric Chown,[a] Lashon B. Booker,[b] and Stephen Kaplan[c]

[a]*Computer Science Department, Bowdoin College, Brunswick, ME 04011;* [b]*The MITRE Corporation, McLean, VA 22102;* [c]*Department of Psychology, Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, MI 48109.* **echown@bowdoin.edu http://www.bowdoin.edu/~echown/ booker@mitre.org skap@umich.edu**

**Abstract:** Perceptual learning mechanisms derived from Hebb's theory of cell assemblies can generate prototypic representations capable of extending the representational power of TEC (Theory of Event Coding) event codes. The extended capability includes categorization that accommodates "family resemblances" and problem solving that uses cognitive maps.

It was almost 60 years ago that the brilliant Cambridge psychologist K. J. W. Craik (1943, p. 61) observed:

> If the organism carries a "small-scale model" of external reality and of its own possible actions within its head, it is able to try out various alternatives, conclude which is the best of them, react to future situations before they arise, utilize the knowledge of past events in dealing with the present and future, and in every way to reach in a much fuller, safer, and more competent manner to the emergencies which face it.

Craik's description is clearly based on the assumption that actions and objects share space in the internal model. Comparably, Tolman's (1948) cognitive map concept implicitly incorporated both action and objects. In fact, MacCorquodale and Meehl (1954) provided a translation of Tolman's concept in their $S_1$-$R_1$-$S_2$ framework. Their formulation of the cognitive map codes the following information: if the organism were to take some action, $R_1$, in the context of a given situation $S_1$, a new situation $S_2$ would result.

This suggests that there is a close kinship between the Theory of Event Coding (TEC) and the world of cognitive maps and internal models. A cognitive map is an internal representation of an environment that is used in much the same way that Craik proposes. Cognitive maps are representations of navigable environments. These representations must also be navigable if they are to be useful as tools for planning or reacting to events. It is difficult to imagine, for example, how one would model a world for navigation if one did not explicitly include a model of movement. Fortunately, one of the simplest, most efficient ways to model experience also naturally encodes actions. The PLAN cognitive mapping model (Chown et al. 1995), for example, stores schematic visual scenes capturing information in a fashion analogous to how it was originally processed. Representations of this type have been called "semantically transparent" by Smolensky (1988) who argues that such representations are efficient because they do not require extra processing to convert them from one form to another. Cognitive mapping models such as PLAN are based on the notion that experience is sequential, with each transition in the sequence arising as the result of an action (locomotion in the case of a cognitive map for navigation). With subsequent experience such sequences are overlayed, forming a kind of network. Planning with such structures, therefore, is a matter of finding an appropriate sequence from start to goal in the network structure. Since the actions are implicitly a part of the representation, they will be naturally extracted as part of the plan. Such structures take the TEC construct a step further by explicitly proposing the form of the internal representation at the cognitive map level.

This kind of network structure derived from experienced sequences of actions is also the basis for the internal representation of the constituent elements of a cognitive map: the representations of percepts, actions, and their relationships. A key issue at this level of representation is perceptual generalization, the process that allows diverse stimulus patterns to be placed in the same category. The event codes proposed in TEC rely on common features and hierarchical abstractions to achieve this important capability. While this approach is a useful beginning, it does not provide all of the necessary functionality. Humans are capable of learning prototypes or "natural categories" in which no fixed subset of features is necessary and sufficient for category membership. In fact, some members in a category may share no common features at all, having only a "family resemblance" as the basis for category membership (Rosch 1978). These observations apply to perceptual categories, action categories, and event categories integrating perceptions and actions. The key to handling the challenges posed by natural categories and family resemblance is to look carefully at the process of perceptual learning. **Hommel et al.** make a key observation along these lines by noting that "perceiving the world is a process of actively acquiring information about the perceiver-environment relationship, including all sorts of movements of eye, hands, feet, and body" (sect. 3.1.1). Perceptual learning mechanisms derived from Hebb's (1949) theory of cell assemblies can generate prototypic representations capable of handling family resemblances (Posner 1973; 1986). Moreover, related control mechanisms ensure that the activation of overlapping or alternative representations will result in a single, definite perceptual outcome (Kaplan et al. 1990; Kinsbourne 1982). For these reasons, the cell assembly framework appears to provide a natural way to extend the representational power of TEC event codes, both intrinsically and as elements of a cognitive map.

One should note that cognitive maps are not limited to navigation in the traditional sense of the word, merely to navigable environments, and it is of course an action space that makes an environment navigable. Any semantically transparent structure with an implicit model of actions will naturally provide a basis for problem solving using the same idea of extracting a sequence from a starting location to a goal location inherent in navigation domains. Models of this type, notably using reinforcement learning (Kaelbling et al. 1996), have begun to supplant more traditional planners in AI systems because generic structures and uniform methodology can be applied and easily learned in virtually any problem solving environment. These observations also apply to cognitive maps and environments that require explicit models of actions. The same mechanisms that generate explicit prototypes for perceptual categories will also produce explicit prototypes for action categories. The same mechanisms that provide for flexible navigation and way-finding can also be used for opportunistic problem solving using explicit action models. In this sense, cognitive maps and the mechanisms that generate and use them are truly generic. The ability to create such generic structures and methods will necessarily be an important step in the development of the TEC model.

ACKNOWLEDGMENTS

The first author gratefully acknowledges the support of the National Science Foundation under grant number IIS-0092605. The second author gratefully acknowledges the support of the MITRE Sponsored Research (MSR) Program.



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

## Common codes for situated interaction

Paul Cisek and John F. Kalaska

*Département de Physiologie, Université de Montréal, Montréal, Québec H3C 3J7 Canada.* **cisekp@magellan.umontreal.ca**
**http://www.cisek.org/pavel** **kalaskaj@physio.umontreal.ca**

**Abstract:** A common code for integrating perceptions and actions was relevant for simple behavioral guidance well before the evolution of cognitive abilities. We review proposals that representation of to-be-produced events played important roles in early behavior, and evidence that the neural mechanisms supporting such rudimentary sensory predictions have been elaborated through evolution to support the cognitive codes addressed by TEC.

The Theory of Event Coding (TEC) is a "framework for the cognitive underpinnings of perception and action planning." In their scholarly introduction, **Hommel et al.** point out that what most traditional approaches to cognition lack is a means of linking perception and action together. TEC proposes to provide this linkage through a common code representing both "perceived" and "to-be-produced" events. A rich body of psychological data is reviewed to support the existence of such a common code.

Like many of the cognitive theories it hopes to replace, however, TEC is developed along what one might call an "inverted phylogeny." That is, it begins by focusing on representational structures necessary for advanced cognitive abilities of humans, such as flexible action planning, and suggests that these abilities require the evolution of cognitive codes for linking perceptions and actions. **Hommel et al.** suggest that common coding "evolved as a means to find a transition directed from responding to given stimulus conditions to planning goal-direction actions" (sect. 5.1), in other words, that it allowed reflex-like behavior to evolve toward goal-directed planning and cognition. However, perception and action have always been closely linked, well before advanced cognitive abilities and cognitive representations ever entered the evolutionary stage. Ancestral animals were quite capable of producing adaptive actions guided by sensory information about their world and aimed at achieving behavioral goals; and abilities such as flexible planning had to have evolved within the context of their situated activity. Thus, the common coding emphasized by TEC most likely reflects the ancestral heritage of cognition itself – the perception-action linkages present for millions of years in the functional architecture for situated interaction.

With this perspective, one is motivated to address the issue of perception-action linkage from a more humble evolutionary viewpoint. How did early organisms guide action? What sensory processing was necessary to organize their activity in the world? A great deal of work has been done on such issues, and many ideas similar to those proposed by TEC have already been discussed in the field. A classic example is the ethological concept of the "Sollwert," or "should-be stimulus" (Hinde 1966). Like event codes, the Sollwert is a representation of desirable stimuli that a behaving system should aim to produce through its behavior. Here, we review some related concepts regarding the control of situated interaction and outline a speculative phylogenetic story of how they may have evolved toward the kinds of cognitive abilities that TEC is meant to address.

The basic functional organization of early living systems was feedback control (for review, see Cisek 1999; Powers 1973). From a general perspective, the processes we call "behavior" are specializations of that control which extend it into the environment by exploiting consistent properties of that environment. In this way, they are related to the processes called "physiology," which maintain internal control by exploiting consistent properties of biochemistry. In order for behavioral control to be effective, actions must be performed that produce consequences that benefit the organism. Early animals could have gotten along with simple trial-and-error behavior and reinforcement learning, but more effective behavior requires the development of neural systems to predict the sensory consequences of potential actions and to use these predictions to select the action that leads to the most desirable results.

Consider an animal situated within a natural environment – for example, a thirsty gazelle approaching a pool of water. Information from the environment specifies to the animal a number of potential actions currently available: the location of the water specifies the direction for an egocentric approach behavior with the possible benefit of quenching thirst, the presence of a lioness near that pool calls for avoidance; and the surrounding terrain, including the layout of obstacles, specifies the possible directions in which to flee. The animal's task is to select those actions whose consequences will have the greatest benefit within its current behavioral context – approaching water is desirable, but only if it doesn't bring the gazelle too close to the lioness. Here, the distinction discussed by **Hommel et al.** between "sensorimotor" and "ideomotor" views is relevant. A sensorimotor architecture is the only means to specify the parameters of available movements because all interaction is constrained by current geometry. Whatever the lofty goals of one's journey may be, one must first successfully navigate through the door of one's house. However, an ideomotor, goal-driven architecture is the best source of information for selecting an action from among the current possibilities. This selection must be biased by both discrete immediate goals such as quenching thirst, and by general requirements such as self-preservation. Such mechanisms had to have existed in at least a crude form even in ancient animals. One might suggest that the simple behavior that laid the foundations of neural organization consisted of first using sensory information to specify currently available potential actions, and then selecting the action best suited toward current goals (Cisek, in press; Kalaska et al. 1998; Toates 1998). Both the selection and guidance of movement requires a means of predicting the sensory consequences of actions.

For example, prediction of sensory consequences is crucial for identifying what is and what is not constant in the world while moving through it. According to the classic "reafference principle" (von Holst & Mittelstaedt 1950), information about self-induced motion such as eye or head movement is used to cancel out the sensory consequences of that motion, leaving a perception of constancy. Without the prediction of sensory events produced by saccades, the world would appear to shift dramatically every time we moved our eyes.

Prediction of sensory consequences is also important for guiding an ongoing movement. For example, even a simple reaching movement cannot be executed quickly by relying solely on actual sensory information, because inevitable conduction and processing delays will render that information outdated. This can result in instability. However, because the laws of biomechanics and physics enforce a tendency for a given movement to produce a consistent sensory result, sensory feedback can be learned and predicted. This predicted feedback can then be used to adjust movement in place of the delayed real feedback. It has been proposed that effective motor control utilizes predictions of sensory feedback, or what in motor control theory is called a "forward model" (Miall & Wolpert 1996). For example, a forward model of arm dynamics can be used to predict the inertial and interaction forces that are generated during a particular movement, and to pre-emptively compensate for these forces, thereby eliminating errors in execution. Our ability to learn new force environments may be due in part to the operation of such a forward model (Lackner & DiZio 1994; Shadmehr & Mussa-Ivaldi 1994).

A forward model can also be used to predict the sensory consequences of entire acts. For example, moving in a particular direction has the reliable consequence of bringing objects in that direction closer. This simple fact can be very valuable for selecting actions. Suppose you see a desirable object that can satisfy some current need when obtained (e.g., a fruit that can satisfy hunger). If the object is within reach, then reaching and grasping actions are available and will be selected because they are expected to result in the greatest payoff. If the object is out of reach, then reaching and grasping actions are not available. However, among other



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

options, it may be possible to approach the object. One predicted consequence of approaching an object is to put it within reach, and to thus make the desired reaching and grasping action available. Thus, a simple sequence of acts such as approaching and obtaining food can be organized using a predictive estimate of how a particular action might make other actions available, and what the expected payoff of those new actions will be.

A growing body of neurophysiological data suggests that the brain is capable of predicting the sensory consequences of its behavior (for review, see Kalaska et al. 1997). For example, some neurons in the lateral intraparietal area (LIP) remap the expected shift in location of a visual stimulus in their receptive fields *before* the initiation of a saccade, anticipating the shift in the visual field that the saccade will produce (Duhamel et al. 1992). When the motion of an object is coupled to an animal's movement, neurons in LIP predict the motion even while the object is temporarily obscured (Eskandar & Assad 1999). The cerebellum has often been implicated in playing a central role in prediction of sensory consequences, appearing to serve the role of a forward model (Albus 1971; Grossberg 1969; Marr 1969; Miall & Wolpert 1996; Rosenblueth et al. 1943).

In light of the discussion above, it is intriguing to consider some aspects of cerebellar anatomy. To summarize it very briefly, the cerebellum forms a set of discrete parallel loops with various parts of the fronto-parietal cerebral network responsible for movement production (Middleton & Strick 2000). Some of these loops pass through the primary motor cortex and are presumably involved in adjusting the descending motor command to compensate for movement dynamics, as discussed above. Other cerebellar loops pass through premotor cortex, and may be involved in predicting the immediate consequences of specific intended movements. The observation of "mirror neurons" in ventral premotor cortex (Fadiga et al. 2000) is consistent with such a sensory prediction hypothesis. Finally, still other cerebellar loops pass through areas 46 and 9 of the prefrontal cortex. These regions are implicated in higher cognitive functions such as sensory integration, working memory, decision-making, and planning (Bechara et al. 1998; Constantinidis et al. 2001; Fuster 2000; Kim & Shadlen 1999; Rowe et al. 2000; Wise & Murray 2000). From the context of the discussion above, they may be involved in integrating current, memorized, and predicted sensory information in order to allow effective selection of action.

One may speculate that the anatomical arrangement of these various cerebro-cerebellar loops is evidence of their phylogenetic relationship. The cerebellar projection that ultimately arrives in primary motor cortex passes through the dorsal portion of the dentate nucleus (one of the cerebellar output nuclei). Just adjacent to that, through the lateral dentate nucleus, passes the projection to premotor cortex. On the basis of their proximity, one may speculate that the projection to premotor cortex evolved through the differentiation and specialization of the already existing projection to primary motor cortex, or that these cerebellar projections may have differentiated in parallel with the expanding cerebrum. Going further ventrally in the dentate, adjacent to the premotor projection, one finds cerebellar projections to prefrontal cortical area 46 and area 9. This again suggests a phylogenetic relationship. Perhaps the higher cognitive functions such as decision-making and planning emerged out of ancestral abilities such as movement control and rudimentary response selection, and continue to make use of the general mechanism for predicting consequences of acts. Perhaps TEC derives from an old architecture for situated interaction which has been predicting consequences of actions for a long time.

In conclusion, many aspects of behavior, including the guidance of ongoing movements and the selection of potential acts, can benefit from mechanisms which combine sensory and motor information to predict the perceptual consequences of actions. Such predictive control probably existed in simple animals and established the fundamental neural organization millions of years before cognitive abilities such as flexible decision-making entered the stage. The control of behavior therefore involved "to-be-produced" events long before cognition – not to link perception and action (which were never separate in the first place), but for simple practical reasons of movement stability and effective behavioral selection. Over the long course of evolution, the kinds of explicit descriptive representations emphasized by most of cognitive science emerged as specializations within the grounded framework of situated activity. They gradually became liberated from context-dependent sensorimotor codes in order to support increasingly arbitrary associations and abstract reasoning. However, as the data reviewed by **Hommel et al.** suggests, they have not been liberated completely. A certain degree of common coding remains in cognitive representations *because* cognition inherited the properties of the functional architecture for situated interaction within which it evolved.

## Unifying by binding: Will binding really bind?

Jörn Diedrichsen[a] and Eliot Hazeltine[b]

[a]*Department of Psychology, University of California, Tolman Hall, Berkeley, CA 94720;* [b]*NASA-Ames Research Center, MS 262-4, Moffett Field, CA 94035.* **jdiedri@socrates.berkeley.edu**
**http://socrates.berkeley.edu/~jdiedri** **ehazeltine@mail.arc.nasa.gov**

**Abstract:** The theory of event coding by **Hommel et al.** proposes that feature binding is a central component of action planning. To evaluate the binding hypothesis, we consider findings from studies of action-perception interference and bimanual movements. We argue that although binding of action features may be a valuable concept, interference from partial feature overlap does not provide a parsimonious account for the observed phenomena.

***Unifying by binding: will binding really bind?*** Progress in the field of physics is manifested in theories becoming more "fundamental." That is, as theories mature, fewer sets of laws are required to understand the workings of the universe. In contrast, within psychology, specialized theories are developed with each new set of data. However, the Theory of Event Coding (TEC) by **Hommel et al.** attempts to unify theories of action and perception, based on the proposal that stimuli and responses share a common representational medium. By virtue of this core tenet, the theory accounts for a broad range of stimulus-response-interaction phenomena, some of which count among the strongest and most reliable in experimental psychology.

One of the most innovative aspects of the TEC is the proposal that feature binding operates not only in perception but also in a similar manner for action. The TEC claims that before an action is executed, its features are bound together into an action plan. Additionally, it is assumed that features can only be bound to a single action plan or object file at a time. Thus, partial feature overlap between the constituent features of two action plans or of an action plan and a perceptual object may lead to interference.

However, if the constituent features are activated but unbound, facilitation may occur. In this way, the notion of binding plays a critical role in predicting patterns of facilitation and interference in the domains of perception and action. We wish to explore whether a unitary binding concept can parsimoniously explain such a diverse and extensive set of findings.

***Action to perception.*** A strength of the TEC is that it provides a framework for interpreting the various findings showing that performing an action can alter the perception of objects. However, any theory attempting to encompass all of these effects faces the difficulty that in similar experiments, partial feature overlap between the action code and the perceptual object leads either to facilitation or to interference.

For example, subjects planning a double key press followed by either a left or right key press are less likely to detect or correctly identify a centrally presented arrow pointing in the same direction



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

as the second key press compared to the situation where the arrow is pointing in the opposite direction. This phenomenon has been called action-effect blindness (Müsseler & Hommel 1997). The blindness for action-congruent stimuli is observed during the planning period of the action before its execution (Wühr & Müsseler, submitted). However, with a different experimental setup, partial overlap facilitates perception. Subjects making a speeded left or right key response to a central stimulus are more accurate in reporting the identity of letter stimuli when their location corresponds to that of the response (Diedrichsen et al. 2000; Hommel & Schneider, in press).

The TEC's account for these results is that in the former case the action features are bound at the time that the arrow is to be encoded, whereas in the latter case, the action features are activated but unbound. However, we doubt that the time course of feature binding is sufficient to explain the opposite outcomes. In both experiments, the stimulus appears before the overt response, and thus the action features should be bound at that time. It is difficult to imagine a time course for the binding that can explain both outcomes.

***Action to action.*** Similar problems arise when binding and partial feature overlap are invoked as the underlying mechanisms for interference phenomena between two actions. The TEC predicts interference between two actions if they share some but not all critical features. This prediction is borne out in a study by Stoet and Hommel (1999), who report that partial feature overlap slows down the production of an intermittent action.

However, this theoretical interpretation seems to be at odds with the well-established finding of bimanual interference. If actions are planned and executed with each hand at the same time, performance is faster and more accurate when they share critical features (Heuer 1993) compared to when the actions differ on all features. For example, reaction times are faster when both hands perform a forward reaching movement than when one hand performs a forward and the other hand performs a sideways movement. This effect is likely due to interference at the level of response-selection (Diedrichsen et al., in press). It is not obvious how the TEC accounts for this finding, because it holds that for partially overlapping actions (e.g., forward-forward), the feature code for "forward" has to be bound into the action plan for both the left and right hand. This competition between the two binding processes for the same code should lead to longer RTs, contrary to what is observed.

Perhaps the two actions are not bound individually but rather form one action plan in the forward-forward condition. If so, we would expect to observe interference from partial feature overlap when an SOA between the first action and the second action is introduced (Hazeltine et al., in preparation; Spijkers et al. 2000). In these studies, participants planned and (sometimes) executed the first movement before knowing the critical features of the second movement. This delay would presumably have required that the two actions be represented as separate entities. Although the TEC's prediction is somewhat ambiguous in this case, it may predict that binding for the second action can start before the binding of the first action is dissolved, especially if the action plans do not overlap. Thus, non-overlapping (i.e., incongruent) actions should be initiated faster than those with greater overlap (i.e., congruent). Although the SOA was varied over a considerable range, no costs for the congruent actions were observed. Again, the idea that partial overlap causes interference seems to find empirical support in only a few experimental setups. For partial overlap to maintain its central theoretical role, interference must be demonstrated when overlap exists between the features of two movements such as direction or extension, and not only overlap in the anatomical side of the effector (Stoet & Hommel 1999).

Although we are concerned that the concept of partial feature overlap may have limited scope, we believe that the idea of binding in the motor domain has the potential to capture the essential nature of the difficulty of performing dissimilar movements simultaneously with two hands. Such situations create an assignment or binding problem in which motor features specifying the movement trajectory must be correctly associated with those specifying the appropriate hand. Indeed, in our bimanual studies (Diedrichsen et al., in press) subjects often confuse the motor plan for the left and the right hands and commit what might be called an "illusory conjunction of action." Thus, the binding of features across perceptual and action domains appears to offer considerable promise in accounting for at least some forms of dual-task interference. However, it remains uncertain if the notion of partial feature overlap will generate much explanatory power for crucial aspects of response selection.

## Modified action as a determinant of adult and age-related sensorimotor integration: Where does it begin?

Hubert R. Dinse

*Institute for Neuroinformatics, Theoretical Biology, Research Group Experimental Neurobiology, Ruhr-University Bochum, ND04, D-44780 Bochum, Germany.* **hubert.dinse@neuroinformatik.ruhr-uni-bochum.de http://www.neuroinformatik.ruhr-uni-bochum.de/ini/PROJECTS/ENB/enb.html**

**Abstract:** Modified action, either artificially induced or occurring naturally during life-span, alters organization and processing of primary somatosensory cortex, thereby serving as a predictor of age-related changes. These findings, together with the interconnectedness between motor-sensory systems and temporally-distributed processing across hierarchical levels, throws into question a sharp division between early perception and cognition, and suggest that composite codes of perception and action might not be limited to higher areas.

The target article by **Hommel et al.** challenges the view that treats perception and action in isolation from each other. Traditionally, brain functioning is regarded as a type of information processing consisting of a chain of sequential events that starts with sensory inputs transmitted on specific pathways to the brain, continues with processing in sensory areas, and culminates in using its outcome to initiate and control action. In a contrary approach, in line with the position advocated by Hommel et al., this chain can be converted by asking what impact action imposes on perception. Using animal experiments we explored how modification of action persistently influences the nature of sensory cortical information processing.

Modified action was induced in rats by unilateral tendon cutting in a hindlimb resulting in moderate, but defined modifications of walking that were reversible about 4 to 5 weeks after manipulation. Electrophysiological inspection of receptive fields, cortical representational maps, and temporal transfer properties of single cortical cells recorded in the hindpaw representation of rat primary somatosensory cortex revealed substantial alterations parallel to those observed behaviorally. Changes consisted of an enlargement of receptive fields, reduction of the size of the cortical hindpaw map, and massive changes in temporal transfer properties (Jürgens & Dinse 1997a; see also Dinse & Merzenich 2001). Cortical alterations developed very rapidly within 2 or 3 days, peaked after 2 weeks, and disappeared parallel to the disappearance of altered walking. The results demonstrate that mild changes of the patterning and style of action substantially modify the way sensory information is processed.

In order to explore whether such an impact of modified action on sensory processing might exist under natural life conditions, we studied sensory processing in old rats. A characteristic age-related modification of the sensorimotor state is a gradual alteration in walking pattern of the hindlimbs (Campbell et al. 1980; Schuurman et al. 1987). We demonstrated that the behavioral alterations were paralleled by massive reorganization of the somatosensory cortex in terms of hindpaw receptive field size and topography of



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

the cortical hindpaw map (Spengler et al. 1995). The apparent correlation between cortical and behavioral alterations makes it unlikely that the observed cortical changes simply reflect age-related degeneration. Instead, we proposed adaptational processes as a consequence of altered walking behavior (Dinse et al. 1995), thereby extending the framework of "modified use" as a determinant of cortical organization to age-related changes (cf. Dinse 2001; Dinse & Merzenich 2001). This view was supported by a number of observations: in case of degeneration, alterations should occur in both the fore- and the hindpaw representations. However, the cortical forepaw representation of animals of even high age revealed no equivalent changes (Dinse 2001; Jürgens & Dinse 1997b). Also, mechanoreceptors of the hindpaw remained largely unaffected by age (Reinke & Dinse 1996). Most importantly, cortical changes in the hindpaw representation could be reversed by either nimodipine ($Ca^{2+}$-blocker) treatment (Jürgens & Dinse 1997b) or by housing the animals under enriched environmental conditions (Churs et al. 1996).

Combined, alteration of action, either artificially induced or occurring naturally during the life span of an individual, alters profoundly the way sensory information is processed. Moreover, the impact of motor-sensory interactions can be used as a predictor of specific features of age-related changes.

What can be inferred from rat data about possible implications for human perception? Plastic changes of rat cortical processing strategies induced by a coactivation learning paradigm resulted in improved spatial discrimination performance in humans subjected to the same coactivation protocol (Godde et al. 1996; 2000), with the gain in discrimination performance correlated with the degree of reorganization observed in primary somatosensory cortex (Pleger et al. 2001). Therefore, extrapolating from our rat data, modification of action will almost certainly impose severe effects on perception in humans.

In a comparative study of the walking behavior of young and old human subjects, severe locomotion impairments became apparent when subjects were studied under high cognitive loads. These results suggest that sensory and motor aspects of behavior are increasingly in need of cognitive control with advancing age (Lindenberger et al. 2000). Accordingly, the walking modification we used as an example of "action" might be underestimated in their relatedness to higher functions and overestimated in their closeness to early sensorimotor processing.

In their article, **Hommel et al.** do not question a functional segregation at early levels, but advocate some forms of shared representations and mechanisms containing composite codes that link late perception and early action. In other words their framework is limited to the cognitive aspects, and implies a division between early sensory-motor processing and cognition.

Anatomical studies in higher mammals agree on the existence of profound cross-talk, feedback, and interactions between motor and sensory information streams. Early cortical sensory and motor areas are heavily interconnected (Neafsey et al. 1986; Scannell et al. 1999). The highest stations of these sensory/motor systems were shown to be associated with prefrontal, cingulate, insular, and parahippocampal cortex and associated thalamic nuclei (Scannell et al. 1999). Given the rich interconnectedness and the massive feedback on the one hand (Felleman & van Essen 1991), and the above described findings from our animal experiments on the other, a limitation of Hommel's framework to a "late" representation that subserves the aspect of a common representation might fail to appreciate the degree of "commonness" present at early stages. It might also underestimate the significance of "early" sensory processing for what might appear "cognitive." In a recent study on Piaget's "A-not-B error," Thelen and coworkers (2001) suggested a dynamical, highly task-dependent organized process that dismisses any type of separation between cognition and motor-perception.

There is another point: processing takes time. There is agreement that in hierarchical systems response latencies increase progressively. However, this alone does not justify the assumption that processing occurs in a step-by-step mode. Response duration must also be taken into account (Dinse 1994; Dinse & Krüger 1994). For the visual pathway, the level-to-level accumulation in latency times is small compared to the overall response durations (Dinse & Krüger 1994). Given an average duration of 50 msec for a neuron response, the time the cell finishes firing the information (i.e., the total time-to-finish) judged by latencies, has reached advanced levels. Excluding the possibility that spikes fired after the first one are meaningless, there can be no sequential processing of the kind, where the proceeding stages wait until the end of the processing of a previous one. What seems more plausible is a continuous interaction through which all stages, including higher-order areas, are linked together in time. This kind of processing is even more evident for other modalities such as the auditory or somatosensory systems that have much shorter response latencies.

Based on these observations, the idea of "temporal distributed processing" was introduced (Dinse 1994), which is grounded on the lack of wait-states: the first spikes carrying information about external events are fed without delay through the entire pathway including high-level areas. Importantly, this routing of information continues until the end of the whole response duration. While those initial response components most likely represent feedforward aspects, different and additional types of interactions and feedback mechanisms come into play as soon as the information is spread through the system. The time required for these processes to have an impact on the response characteristics at a given site might be variable, depending on the type of stimulation and modality. In any case, temporally delayed influences arising from interactions modify the informational content being processed and represented. As a result, even low levels carry signatures of advanced processing steps, such as those related to a common representational medium containing composite codes of perception and action. These different lines of arguments, taken together, make it unlikely that common forms of motor-perception representations are limited to higher areas.

ACKNOWLEDGMENTS

The author acknowledges the support of the Institute for Neuroinformatics, Ruhr-University, Bochum. Part of this work was supported by grants from the Deutsche Forschungsgemeinschaft (Di 334/10-1,2,3 and 334/15-1). I am grateful for insightful discussions with Drs. Marianne Jürgens, Heinrich Reinke, and Ben Godde, and my colleagues at the Neuroinformatics Institute.

## Event coding as feature guessing: The lessons of the motor theory of speech perception

Bruno Galantucci,[a] Carol A. Fowler,[b] and M. T. Turvey[a]

[a]*University of CT, Box U-1020, Storrs, CT 06269;* [b]*Haskins Laboratories, New Haven, CT 06511.* **bruno.galantucci@uconn.edu carol.fowler@haskins.yale.edu turvey@uconnvm.uconn.edu**

**Abstract**. The claim that perception and action are commonly coded because they are indistinguishable at the distal level is crucial for theories of cognition. However, the consequences of this claim run deep, and the Theory of Event Coding (TEC) is not up to the challenge it poses. We illustrate why through a brief review of the evidence that led to the motor theory of speech perception.

We agree with **Hommel et al.** that unification of theories of perception and action is necessary to understand cognition. Equipping animals with efficient perceptual capabilities is expensive for nature. The enterprise would make no sense unless it was motivated by equivalent gains in action capabilities. Conversely, complex elaborations of the action capabilities of an animal would be pointless if not accompanied by perceptual systems that support them. Evolution, unlike modern cognitive science, cannot loosen



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

the functional connections among its products. Thus, perception and action should not be studied in isolation; most likely perceiving-acting is not a decomposable system.

The main assumption of the target article is a logical corollary of this broader claim and, as such, we agree with it. However, the Theory of Event Coding (TEC) does not fulfill its promise, and it makes the very mistake it is meant to correct.

We will focus on the authors' brief, dismissive comments about the motor theory of speech perception. The authors provide two reasons for their dismissal. First, they do not wish to claim that their account applies to language. Second, in the opinions of Jusczyk (1986) and Levelt (1989), empirical support for the motor theory is weak. We address what the authors prefer not to because we consider the reasons for their dismissal mistaken.

However, we will not reply to the second reason here. Evidence from half a century of research cannot be summarized as tersely as a dismissal can be offered (for a review, see Liberman 1996). As for the first reason, speech is an exquisite example of perceptually guided action, and thus it must be addressed by a theory of perception-action codings. Moreover, precisely because it cannot be reduced to the stimulus/response experimental settings the authors choose to use, it is the right place to look for a better understanding of the couplings between perception and action. And the empirical facts that led to the motor theory reveal the flaws of TEC to which we have alluded.

In the tasks that support TEC, experimenters devise stimuli that can be described by sets of arbitrarily-chosen, arbitrarily-combined features (e.g., a letter is red or green; a rectangle is on the right or left side of a computer screen). In the tasks, some of the features are relevant, and so the participants are encouraged to make use of them to do the task, and they do. However, these sets are not up to the task of constituting percepts or action plans in nature. Proponents of TEC have to answer, among others, the following questions:

1) Can such feature sets compose real perception-action codings?

2) Are percepts linear combinations of features?

3) If they are to refer to the distal world, perception-action codings must be grounded. How are features grounded?

The motor theory of speech perception was designed to answer these questions, which arose when Alvin Liberman and colleagues sought the features that capture our perceptions of syllables (Liberman 1957; Liberman et al. 1967; Liberman & Mattingly 1989).

They started where Hommel and colleagues would have liked them to start. During the fifties, as a part of a project to build a reading machine for the blind, they tried several ways to teach people to perceive syllables as strings of featurally distinct sounds. They devised "acoustic alphabets": acoustic features corresponded to single letters, and syllables were sequences of discrete acoustic letters. After years of effort, the project failed. Speech is not an acoustic alphabet, and people could not perceive acoustic alphabetic sequences at practically useful rates.

This realization provides an answer to our first question: arbitrarily-chosen, arbitrarily-combined features do not correspond to perception-action codings.

The next move of Liberman and colleagues was also aligned with TEC. Perhaps the sounds created by the vocal tract have a special character for which arbitrary sounds cannot substitute. Using the sound spectrograph, Liberman and colleagues searched for acoustic features that, they hoped, would characterize syllables. Again, their hopes were frustrated. Spectrograms, far from clarifying the picture, presented a new puzzle. Depending on the context, the same acoustic feature could specify different phonemes and, conversely, different acoustic features could specify the same phoneme. Acoustic features do not capture the dimensions of the space where syllables live.

We can now answer our second question: percepts are not necessarily linear combinations of acoustic features. But, more importantly, this second failure taught Liberman and colleagues a lesson. Once they rejected features as components of speech percepts, they recognized that there is an invariant among different tokens of the same phoneme; it is in the gestures that produced them. This discovery led to a major revision of their scientific assumptions. Motor competence, not the feature set conceived by the scientist, underlies speech perception. The motor theory was born, and it provided a surprising answer to our third question: speech perception is not grounded in the realm of perception, but in the very place where, according to the common coding of distal events, it should be grounded: in the actions of the vocal tract!

To understand the nature of perception-action codings, a major revision of the conceptual apparatus used to implement TEC is needed, beginning with careful decisions about what should and should not count as an observable. Liberman, and with him Roger Sperry (1952), suggested that the use of two different sets of observables – one for perception and one for action – is misleading; motoric observables capture cognition. Although we do not entirely accept this full motorization of cognition, the hard learned lessons that led to the motor theory tell us that the observables conceived by a theory of disembodied cognition will not work. Here lies the major limitation of TEC. It rejects the unreasonable division between action and perception at the level of distal events, while claiming that "late action" and "early perception" are not necessary to explain cognitive common coding. If not there, among the nuts and bolts of the contact between the nervous system and the physical world, where are we to find the grounding of perception-action distal identity?

## A common framework for language comprehension and language production?

Robert J. Hartsuiker and Martin J. Pickering

*Department of Psychology, University of Edinburgh, Edinburgh, EH8 9JZ United Kingdom.* **Rob.Hartsuiker@ed.ac.uk**
**http://www.psy.ed.ac.uk/~robh/** **Martin.Pickering@ed.ac.uk**

**Abstract:** Natural language processing involves a tight coupling between action (the production of language) and perception (the comprehension of language). We argue that similar theoretical principles apply to language processing as to action/perception in general. Language production is not driven solely by the speaker's intentions; language comprehension is not only input-driven; production and perception use common representations. We will relate recent findings from our language production lab to the Theory of Event Coding (TEC)'s principle of feature binding.

**Hommel et al**. consider standard approaches studying action and perception in isolation to be flawed. They argue that action does not start solely from an organism's goals and intentions, nor is perception input-driven only. The authors present what they refer to as a set of "metatheoretical principles" (TEC) in which action and perception are mutually dependant processes, sharing a common representational medium (event codes). In this commentary, we will argue that these metatheoretical principles can be and should be extended to the field of natural language processing (psycholinguistics), a field of study that the authors explicitly exclude from their target article (see sect. 2.1.4.1). However, in our opinion, language should not necessarily fall outside the scope of their *theory*, for two reasons. First, language processing in natural dialogue consists of a cycle of tightly coupled events, that consecutively involve action (producing an utterance) and perception (perceiving the interlocutor's utterance). These events are closely related. However, like older research traditions in action or perception, research in language has mainly considered production and comprehension as isolated events, ignoring this tight coupling. Second, **Hommel et al.** argue for a common representational medium for action and perception. We argue that the same is true for linguistic acting and perceiving.

***1. The fundamental concern of psycholinguistics should be***



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

***dialogue, not monologue.*** Current research on language processing tends to view production and comprehension as isolated events. Views on language production can perhaps best be characterized by the subtitle of Levelt's (1989) influential book, *Speaking from intention to articulation,* which clearly reveals the essential ideomotor view that presently dominates production research. Likewise, research on language comprehension mostly studies the reading of isolated words or sentences. This line of research attempts to exclude responses (perhaps most successfully in ERP studies on sentence comprehension, where the lack of the need for a response is regarded as a virtue) or to ignore them (e.g., by tacitly assuming that word naming latencies reflect only perceptual aspects of reading; see Bock 1996). However, such research does not do justice to the fact that language is usually produced in dialogue, where perception/production interactions are of central importance (Pickering & Garrod, submitted). Consider a question-answer pair. It is not only the goals and intentions of the answer-giver that drive the content and form of the utterance: it is also the content and form of the question that is being asked. There are also perhaps less obvious phenomena in dialogue that suggest a tight coupling between production and perception. Almost any statement calls for a particular type of response from the interlocutor; dialogue partners often co-produce utterances (you start a sentence and I finish it); dialogue partners tend to "align" their representations (i.e., they tend to use the same linguistic units at each level of processing); they monitor and correct not only their own speech but also that of the other. Any theory of language use needs to account for this coupling between production and comprehension processes.

***2. Common representations for production and perception.*** A fundamental assumption of TEC is that action and perception make use of common codes. Similarly, there is growing consensus in psycholinguistics that common representations are used in language production and language comprehension, although theorists are still arguing whether this is only true for abstract lexical representations (Levelt et al. 1999), or for most linguistic levels (MacKay 1987; Pickering & Garrod, submitted). One piece of evidence for this parity of representations is the finding of syntactic priming from comprehension to production (Branigan et al. 2000). Producing a particular syntactic form (e.g., a passive) increases the likelihood of that form being produced again on a subsequent trial (Bock 1986). However, Branigan et al. showed that *perceiving* a particular sentence type is sufficient to create a priming effect in production, suggesting a common code for perceptual and production codes for syntax. According to Pickering and Garrod, common representations play a crucial role in dialogue, because they allow a pair of speakers to achieve "alignment" (similar representations at each level). This alignment is their main theoretical construct for explaining the fluency of dialogue.

***3. Effects of feature overlap in language processing.*** One example of the possible viability of TEC as an explanatory mechanism for linguistic processes comes from the domain of syntactic priming. There is now a large body of evidence from our group that this effect is modulated by the degree of lexical overlap (e.g., Pickering & Branigan 1998): the more words are shared between prime and target, the more syntactic repetition. Why is this? From a TEC perspective, one could consider a sentence an integrated representation binding syntactic information and lexical information. If a particular structure, say a prepositional object dative, is bound to particular words (e.g., *the nun showing an apple to a pirate*) and if a subsequent picture shows one or more of these elements (say, the nun giving an apple to a priest), then it would be relatively costly to say *the nun handing the priest the apple* because it would require undoing the binding between particular words and a particular phrasal configuration. The more words are shared, the costlier the undoing of these bindings would become.

***4. Conclusion.*** Current-day psycholinguistics is still too much focussed on sentence-production and comprehension in isolated contexts (monologue) and tends to ignore action/perception interactions, except in some circumstances such as self-monitoring of speech (Hartsuiker & Kolk 2001; Levelt 1989) or in some approaches that stress the role for production in comprehension (Garrett 2000). We think that psycholinguistics would benefit from the controlled study of language in dialogue, casting findings in theoretical frameworks such as TEC.

## TEC – Some problems and some prospects

Julian Hochberg

*Psychology Department, Columbia University, New York, NY 10027.*
**hochberg@columbia.edu**

**Abstract:** The Theory of Event Coding (TEC) is a significant contribution to the study of purposeful perceptual behavior, and can be made more so by recognizing a major context (the work of Tolman, Liberman, Neisser); some significant problems (tightening predictions and defining distal stimuli); and an extremely important area of potential application (ongoing anticipation and perceptual inquiry, as in reading and movies).

**Hommel et al.**'s paper on the Theory of Event Coding (TEC) is a significant contribution toward a science of purposeful perception and action, one in which the course of intended actions, like key-pressing, is determined by component action codes that share a common cognitive representation with stimulus feature codes ("red," "round," etc.). Supporting research mainly concerns interactions between stimulus and response features as associatively formed during instructed repetitive practice.The authors stress that the stimuli that most directly determine the event are not proximal but distal, by which they mean the perceived features in represented space. This, they rightly contrast with the common approach in which stimuli and reponses are defined in terms of the physically measurable proximal stimulus input and effector output. (This particular use of the word "distal" may prove momentarily confusing to perception psychologists, since it is not a 2D projection vs. 3D layout that is at issue.)

Their approach is not quite as uncommon or isolated as their presentation suggests. Tolman's historic analyses (1932; 1948) of animals' purposive behaviors in, and memories of, their behavioral environments in response to the patterns that offer a "means-end readiness" (since termed an "affordance" by Gibson 1979; see also Lewin's valence-driven events in life space), were instantiated by specific experimental demonstrations, like place learning, in which behavior was far better predicted by measurable variables in distal space than by proximal stimulus measures. This approach was not abandoned, nor is it rare. Research on place learning and its purpose-relevant stimuli is not obsolete (cf. Gallistel 1990), and research into the brain processes of topological learning is cutting edge (cf. Zinyuk et al. 2000). Indeed, closely related "constructivist" approaches in which behavior is explained in terms of events that transpire in an internalized space abound today, although not in connection with Reaction Time (RT) experiments, and with little work until now on the explicit relationship between the goal, the stimulus information, and the action.

What *is* uncommon about TEC is that it appears to set the stage for systematic co-analyses of such purposive behavior, in which the signal and the action are closely related and commensurate. There are, however, several serious problems to be resolved and opportunities grasped, before we can be as comfortable as we might be with TEC as a framework. Namely, in ascending order of difficulty:

(1) More explicit modeling is needed than has been reported in the target article, in order to reach conclusions that are more definite than the ones the paper typically offers, for example, that such-and-such an outcome is not incompatible with TEC.

(2) Identifying and codifying of feature codes is needed if they are to be applied by more than intuition. The "distal stimulus" needs principled, circumstance-dependent definitions: the authors obviously do *not* mean measures defined only in terms of the



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

physical measures on the distal environment, since they recognize that the effective stimuli are motivation dependent. This means that some standardizable methods of determining what the stimuli (or their "distal features") are, have to be established.

(3) Perhaps the greatest problem, as well as the greatest opportunity for this approach, and for psychology as a whole, is to step beyond the authors' almost-paradigmatic RT experiment to the purposeful, self-guided sequential interaction of stimulus and response that prevail in our normal perception-guided/perception-guiding behavior. There are indeed many behaviors that we need to examine from something like a TEC position – and to bring under real scientific inquiry – because they are a large part of our life. They require appropriately designed presentation conditions because, even if they do not eventually connect with a bottoms up approach – and they may not – they are loaded with practical applications (in engineering, in medicine, in education, and in entertainment).

Specifically, I hope and expect that TEC will advance into the research and analysis that the communications media now make possible and, indeed, demand, as text and moving pictures are increasingly created by computational devices. What you perceive in the individual glimpses provided by text or movies, and carry forward through subsequent glances, is of course determined in part by low-level early factors; more important and significant, however, are the anticipations and questions that are formed – the perceptual purposes that are pursued – and the eye movements that carry and test the process.

The perceptual inquiry involved in skilled reading, as in listening to speech, is not a passive receptor process. Rather, it is surely an active planning and testing process, without which the stimuli (distal or otherwise) have no meaning. Liberman et al. (1967) proposed that listening proceeds via an analysis-by-synthesis – by active anticipation and testing. **Hommel et al.** shy away from Liberman for reasons which may fit their particular presentation, but which overlook the avenues opened by his approach: A very similar view of active looking in both reading and in visual attention and consciousness (Hochberg,1968; 1970; 1997; O'Regan & Noë 2001), closely related to Neisser's seminal work (1967), seems ready to profit from something like TEC. This approach has been extended to the behavior involved in actively using the information in movies (e.g., Hochberg & Brooks 1996). Of course, not all reading or listening or movie-watching is of the same kind or level of inquiry, and short-term (low-level) effects intrude continually (e.g., in "bad cuts"), but these can be fairly readily systematized by the TEC-like knowledge we need.

## Attending, intending, and the importance of task settings

Jason Ivanoff and Raymond Klein

*Department of Psychology, Dalhousie University, Halifax, Nova Scotia, Canada B3H 4J1.* **ji@or.psychology.dal.ca** **ray.klein@dal.ca** **http://www.dal.ca/~klein/**

**Abstract: Hommel et al.** emphasize that the Theory of Event Coding (TEC)'s utility is not its ability to be a new theory of cognition, but its ability to engender new thinking about new and old problems. In this commentary we use the TEC to re-examine a long-standing discrepancy in the attention literature.

The classic theories of attention, as applied to linear stage models, pose that it functions to completely (Broadbent 1958) or partially (Treisman 1960) filter irrelevant early information, or increase the relative activation (LaBerge & Samuels 1974) of relevant features or objects in the environment. Likewise, intention involves preparation for certain features of a movement (Rosenbaum 1980). Although classical linear stage theories emphasize a dissociation between attention and intention because perception and action were viewed as relatively distinct, the notion of a response selection stage that is responsible for translating codes from perception to action provides an arena for interactions between these isolable subsystems.

**Hommel et al.**'s Theory of Event Coding (TEC) theory differs from classic linear theories in that high-level perception and action merge (rather than translate) at the event file (Hommel 1998; or object file; Kahneman et al. 1992). The circular relation between perception and action in TEC is not unlike the TOTE feedback loop in *Plans and the Structure of Behavior* (Miller et al. 1960). The TEC proposes that attention involves modifying the weight between stimulus codes and feature codes in the event file, and intention involves modifying the weights between feature codes and response codes. This idea is not unlike the classical linear stage theory, and therefore it is not particularly novel. What is interesting, however, is the important role of task settings (e.g., as in Norman's concept of pertinence; Norman 1969, pp. 32–36). Features that are most important to the completion of the task are weighted more strongly than others are. In other words, the task-relevant features receive differential processing because, otherwise, the task could never be implemented.

In attentional research, there has been growing interest in the influence of task settings. For instance, Folk et al. (1992) discovered that the performance effects of exogenous orienting to sudden-onset or colour-defined peripheral cues is only evident if the target is defined by sudden-onset or colour, respectively. This finding suggests that exogenous orienting is not fully automatic but partially contingent on task demands.

As the authors proclaimed, TEC is not a theory that is readily falsifiable, but its value lies in its applicability to current issues. A hotly debated theory in the attentional field concerns the premotor theory of attention (PMTA; Rizzolatti et al. 1987), and its predecessor the oculomotor readiness hypothesis (OMRH; Klein 1980), which claims that spatial shifts of attention are the result of the activation of eye movement (Klein 1980) or other motor programs (Sheliga et al. 1994). This idea not only clashes with classic theories that make a clear distinction between perception and action, but it also clashes with TEC's claim that attending involves modifying the weights *to* the feature codes in the event code and intending involves modifying the weights *from* the feature codes to response codes. Although attending and intending are similar processes (resulting from the modification of weights), they occur at different processing stages (i.e., before and after the event code, respectively). OMRH-PMTA, on the other hand, maintains that attending and intending are the result of the same underlying mechanism (e.g., attention is accomplished by preparing a movement).

The empirical evidence has been mixed with some reports that clearly do not support the OMRH-PMTA (Klein 1980; Klein & Pontefract 1994), some reports that do support it (e.g., Deubel & Schneider 1996; Deubel et al. 1998; Sheliga et al. 1994), and some that is completely ambiguous (Corbetta et al. 1998). If one considers that perception and action processes converge in a common arena (the event code), then this apparently mixed field of findings may not be so incomprehensible. The important factor to consider is that "events" are formed on the basis of task conditions. Thus, by considering task settings, TEC may help to explain the apparently discrepant findings

In dual-task studies providing evidence in favour of the OMRH-PMTA – such as when attending appears to influence eye movements (Sheliga et al. 1994) or the advantage of identifying stimuli at a location where an eye movement (Deubel & Schneider 1996) or hand movement (Deubel et al. 1998) is prepared – the tasks (motor and attentional) are frequently learned simultaneously. The implications of this methodology, in terms of TEC, is that attentional processes and intentional processes may be assimilated into a common event code, because of task-settings. On the other hand, the studies that have not supported OMRH-PMTA (Klein 1980; Klein & Pontefract 1994) have used techniques where each "task" is learned in isolation before being combined (i.e., to de-



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

termine whether programming an eye movement improves perception and whether attending to a location gives rise to oculomotor benefits). This means that each task evolved in its own relative task-space and, as such, there would have been little, if any, chance for any interactions.

There is some support for the idea that the relationship between intention and attention is mediated by task-settings. Fischer (1999), using the dual-task methodology, recorded reaction times to probe targets presented to the left or to the right while participants were reading text, scanned pseudo-text as though they were reading, or searched for a target amongst distractors in a left-to-right serial manner. If preparing an eye movement also implicates the allocation of attention to the same location, then probe RT ought to be fastest when the probe stimulus appears to the right of fixation. The result showed that this was only the case in the search task. Apparently, when the two tasks share a common task component (i.e., search), then there is an interaction between the tasks because, in the language of TEC, there is common coding between the tasks. When the tasks were distinct (e.g., reading and search), there was less opportunity for common coding, and the result was no influence of eye movements on detection.

Clearly this reinterpretation will require more specific, and direct, tests. The important point here is that by utilizing a specific aspect of the TEC, a long-standing debate can be refreshed and re-examined. Hence, the TEC may prove to be conceptually useful, at least in this one context.

## The Theory of Event Coding (TEC)'s framework may leave perception out of the picture

J. Scott Jordan

*Department of Psychology, Illinois State University, Normal, IL 61790-4620.*

**Abstract: Hommel et al.** propose that action planning and perception utilize common resources. This implies perception should have intention-relative content. Data supporting this implication are presented. These findings challenge the notion of perception as "seeing." An alternative is suggested (i.e., perception as distal control) that may provide a means of integrating representational and ecological approaches to the study of organism-environment coordination.

**Hommel et al.** present an impressive body of empirical work in support of their meta-theoretical notion that action planning and perception share common mediation. The theory may be more successful than initially anticipated, however, for if it truly is the case that perception and action planning share common resources, we may have to re-evaluate what we mean when we use the word perception. From a folk-psychological perspective, to perceive means to "see." According to the notion of common coding, however, what we see shares resources with what we intend. This implies that perceptions should entail intentional-relative content. Data indicate this to be the case. Kerzel et al. (in press), for example, report that although the vanishing point of a moving stimulus tends to be perceived at a location beyond the actual vanishing point, the displacement reduces to zero if subjects are not allowed to track the moving stimulus via smooth-pursuit (i.e., they are instructed to fixate a different, stationary stimulus during the presentation of the moving stimulus). Jordan et al. (2002) recently discovered that this difference has more to do with action planning than effector motion.

Specifically, they asked subjects to fixate a stationary stimulus in all conditions, yet varied the intentional stance (i.e., action plan) subjects were to maintain in relation to the moving stimulus. In the "induction" condition, the computer produced the offset of the moving stimulus, while in the "intention" condition, the subjects produced the offset of the stimulus via a button press. Despite the lack of eye-movements in either condition, perceived location was further in the direction of stimulus motion in the "intention" condition than in the "induction" condition. To determine whether this difference was due to the actual button press or the action plan, Jordan et al. replicated the "intention" condition along with a "cue" condition in which subjects were asked to press the button in response to stimulus onset, while stimulus offset was produced by the computer. Despite the fact that both conditions entailed similar effector movements (i.e., pressing a button), perceived vanishing points were further in the direction of stimulus motion in the "intention" versus the "cue" condition.

Collectively, these data imply that the degree to which one is ahead or behind in perceptual space varies with the relationship one is specifying and attempting to maintain with the environment. This, of course, is consistent with **Hommel et al.**'s notion of common coding. The existence of such intention-relative perceptual content, however, further implies that perception is not as much an act of "seeing" as it is an act of controlling distal effects. Both the specified distal effect (i.e., the intention) and the current environmental state (i.e., the feedback) are present within perceptual content. Seen in this light, the Theory of Common Coding (TEC) runs the risk of eliminating the descriptive utility of one of the very concepts it strives to explain (i.e., perception). Perhaps we should take this as an achievement of the TEC and give up the concept "perception" or, at the very least, give up on the notion that perception constitutes "seeing." If we were to replace the concept "perception" with the concept "distal effect control," we might actually find that the representational and ecological approaches are much more compatible than initially imagined. Both have always been concerned with how it is that organisms offset environmental perturbation in order to maintain organism-environment coordinations, and both have also agreed that information constitutes the legal tender of such control. Given the intention-laden nature of perceptual content revealed by Common Coding, it may also be possible for the two to agree as to the content of the controlled information. Instead of perceptual content being defined as information about the external world, Common Coding reveals it to be information about "being in the world." Any student of Gibson should quickly recognize the similarities between this idea and the concepts: "optic flow" and "affordance."

## Ecological information and prospective control without mental representation

Nam-Gyoon Kim[a] and Judith A. Effken[b]

[a]*Department of Psychology, William Paterson University, Wayne, NJ 07470;* [b]*College of Nursing, University of Arizona, Tucson, AZ 85721-0203.*

**kimn@wpunj.edu** **jeffken@nursing.arizona.edu**

**Abstract:** We agree with the authors that an adequate account of the perception-action interface is needed. However, we believe that the answer will not be found in a common representational structure encoding distal events, as the authors propose, but in Gibson's notion of ecological information, which, as we demonstrate, specifies not only perspective but also prospective and retrospective states of affairs.

As the target article by **Hommel et al.** points out, perception and action planning typically are considered to be discrete, isolated processes. Perception is assumed to create and maintain a general-purpose, three-dimensional scene description that is independent of the behavior it serves (e.g., Marr 1982). Hommel et al. propose an alternative framework that recognizes the tight coupling between perception and action planning. In their framework, a common representational structure encodes distal features (i.e., events and actions in the animal's environment), rather than proximal sensory data or muscular regularities.

We agree that traditional approaches are inadequate to address



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

the perception-action interface. However, we disagree with the authors' claim that an answer will be found by postulating a common representational structure for perception and action planning. In fact, we are quite dubious about the authors' contention that "Through integrating information from multiple sources . . . distal feature codes can become more complex than proximal codes . . . as complex as 'sit-on-ableness' to take one of the standard 'affordances' of Gibsonian approaches" (sect. 3.2.1.). Unless the authors subscribe to naive realism, this assumption is too optimistic.

The authors seem to have adopted the stimulus-response connection as their model for the perceiving-acting relation. In our view, this model is fatally flawed. Rarely is action interpretable as responses elicited by stimuli or in the presence of stimuli. Neither is the perceiving-acting relation interpretable as a unidirectional effect of one upon the other. We believe, following Gibson, that the ease and accuracy with which animals get along in the environment can only be understood by recognizing that animal behavior is controlled by information, rather than being a series of triggered or emitted responses to stimuli.

Gibson (1950; 1966; 1979) devoted his entire career to this problem. His effort culminated in several new concepts (e.g., the optic array, affordances, and exploratory – as opposed to performatory – activity). These concepts formed the basis for a new theory in which perception is not simply a response to the physical world or a cognitive interpretation of the effects of the world on us, but an active, ongoing process that enables us to gain a genuine awareness of the meaningful, goal-relevant environment. Ultimately, Gibson viewed behavior as being neither physically caused (i.e., the sensorimotor view of action) nor mentally caused (i.e., the ideomotor view of action), which led to his rejecting not only behaviorism, but also the causal theory of perception.

**Hommel et al.** pay tribute to Gibson for advancing interactions between perception and action but denounce his rejection of mental representation as "anathema to ecological approaches." The basis for this denunciation may be their narrow interpretation of the concept of ecological information. A key issue motivating the authors to resort to cognitive representations of a common event code seems to be the future-tending, anticipatory nature of goal-directed movements, that is, movements directed at "what must be done" instead of "what is." With an incomplete understanding of ecological information, accounting for future tending, goal-directed behavior can be a daunting task.

Like the authors, the Soviet physiologist Bernstein (1967) assumed that perception provides only awareness of "what is" but not an awareness of "what must be done" and consequently cannot resolve "the problem of action . . . the reflection or model of future requirements" (p. 171). This led Bernstein to resort to probabilistic extrapolation as the basis for modeling the future. However, this is a highly questionable strategy. Patently, experience with a given situation and its contingent events can lead to an expectation of what is likely to occur when the situation reoccurs, and such expectations can shape the selection and conduct of actions. But an *expectation of what might occur* is not the same as *specification of what will occur,* and the latter is necessary for successful prospective control.

How can information specify, not only the present, but also the future, thereby providing sufficient constraint for the prospective control of movements? As a paradigmatic case, consider long-jumpers who must regulate their final few strides to the take-off board while maintaining maximum controllable velocity (Lee et al. 1982). Careful observation suggests that step regulation during the final zeroing-in phase is not a preprogrammed act, as in a single ballistic adjustment, but the result of continuous coupling between perception and action. Although alternative explanations are possible, we believe that the most parsimonious is a recursive function within the information source itself. In recursive functions, each value entails the next, thus making it possible to convert the present into the future (Rosen 1991).

Extending these insights, we used Taylor's Theorem to generate a recursive function of *tau* (the optical information specifying time-to-contact between an observer and an upcoming surface) and demonstrated that future states of *tau* can be specified given current values of *tau* and *taudot* (the rate of change in *tau;* see Kim et al. 1993; 1998; and Lee 1976, for further details). In tasks that demand planning actions several steps ahead of time, *taudot* and *tau* together comprise the minimal recursive function which can be utilized by actors to access future values of *tau*s so that they can modify their behavior ahead of time, thereby achieving precise regulation – an example of prospective control (Kim & Effken 1995; Kim & Turvey 1998). See also Kim et al. (1995) for cases demonstrating prospective (future-tending), perspective (present-tending), and retrospective (past-tending) aspects of information.

In sum, information in Gibson's specificational sense provides the basis for prospective control (e.g., Turvey 1992) that **Hommel et al.** recognize is necessary. Prospective control allows actors to modify an approach on the fly, as goal-related information and situation demands dictate. Moreover, prospective control does not require the complex, integrated internal representation of perception and action the authors propose. Instead, the answer lies in the specificational nature of perceptual information, further constrained by the goals and effectivities of the observer/actor. In such a mutually constrained dynamical system, no "mysterious or arcane explanatory devices" (O'Regan & Noe, in press), such as those proposed by the authors, are needed to link perception and action planning.

## Exploring the hyphen in ideo-motor action

Wilfried Kunde

*Department of Psychology, Roentgenring 11, 97070 Wuerzburg, Germany.*
**kunde@psychologie.uni-wuerzburg.de**

**Abstract:** This commentary focuses on **Hommel et al.**'s inferences on action planning. It discusses the relevance of anticipated extrinsic movement effects for action control, the problems of a feature-based representation of actions, and the necessity of the acquisition of conditional movement-effect associations.

The Theory of Event Coding (TEC)'s core assumption that actions are represented in terms of – and thus accessed through – their sensorial effects dates back to William James's (1891) ideo-motor theory. In a scathing criticism, Miller et al. (1960) concluded that ideo-motor theory had done little more to bridge the gap between ideas and behavior than inserting a hyphen between them. This may apply to James's introspective approach. Yet, **Hommel et al.**'s excellent target article demonstrates that the ideo-motor gap has turned from a "miracle" into a topic of rigorous empirical investigation. Still, we are exploring the "hyphen," and I want to discuss some shortcomings of recent research, suggesting necessary extensions for the future.

***Pertinence of anticipated distal effects to the study of action control.*** Evidence for the impact of action effects on behavior stems primarily from experimental situations where (1) action effects are presented prior to action execution, or (2) action effects consist of merely movement-intrinsic feedback. Such conditions, however, are less representative for goal-oriented action which by definition is controlled by forthcoming, and thus necessarily anticipated, distal effects. Research relying on action induction and on movement-intrinsic features may therefore limit possible insights into the underlying mechanisms of goal-oriented action.

The fact that the perception of action effects (or effect-resembling stimuli) can induce motor-patterns that typically produce these effects is certainly interesting for a perception-action theory (cf. target article, sect. 2.3.2.2). Yet, this finding does not allow us to infer that effect codes become activated in cases where they are not externally stimulated, nor does it explain how this endogenous activation (i.e., anticipation) of effect codes is accomplished. It is



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

very likely that constraints of action control – as apparent in the psychological refractory period effect or dual-task costs – originate from limitations in the *endogenous* generation and maintenance of effect codes (Greenwald & Shulman 1973). However, action induction provides no appropriate inferential tool to explore these constraints.

Studies dealing with intrinsic (mostly spatial) action features are of limited significance, not only because moving for the sake of the movement is an exception, but because such features are mingled with inherent properties of the motor apparatus that brings the action about. For example, interference between two ipsilateral body movements may result from overlap of spatial action features (Stoet & Hommel 1999). But it may also occur because concurrent ipsilateral movements produce an unstable body posture, or because ipsilateral movements violate basal biological motor-patterns that predominantly comprise activity of contralateral limbs (walking, crawling, etc.), or because they are processed in the same brain hemisphere. Indeed, when actions are coupled orthogonally with extrinsic effects, no effect-overlap costs (rather than overlap benefits) emerge (Kunde et al., in press).

Thus, a thorough understanding of voluntary action will likely require a focus on the impact of action effects that are, (1) internally anticipated (i.e., follow but do not precede the action) and are, (2) extrinsic (i.e., occur outside the body). Although **Hommel et al.** acknowledge this requirement (cf. sects. 4.2.2 & 5.1) there are few studies that fulfill *both* criteria so far, which leaves an uncomfortable discrepancy between TEC's theoretical scope and its current experimental reality (cf. Kunde 2001; Steininger 1999).

***How to transform an abstract, feature-based action plan into a concrete motor pattern.*** TEC suggests that action planning, analogously to (visual) perception, relies on elementary features. In my opinion this analogy is questionable. Whereas the perceptual *analysis* of objects can often be limited to very few attributes (in pop-out search even to a single feature), the *synthesis* of a to-be-performed action cannot be so limited. Features like LEFT and HAND might somehow constrain motor-activity to the left hand, but they fall short of determining one particular out of the plethora of possible left-hand movements. This problem becomes even more apparent when considering more extrinsic effects. Pressing a light switch can be achieved by motor patterns as different as lifting an arm or swinging a hip. How is the inevitable specification of one particular out of the infinite number of equifunctional motor patterns achieved? To me it appears an unappealing theoretical gap that TEC delegates this task to an unspecified "machinery of the 'late' motor processes" (target article, introduction). To become a comprehensive theory of action control, TEC should incorporate, rather than keep out, mechanisms that select one particular out of the usually large number of goal-satisfying behavioural options.

***Inconsistent evidence for feature-binding in action planning.*** **Hommel et al.** assume that features needed to represent a certain event are temporarily less available for the concurrent representation of other events, producing costs of feature overlap. TEC has done an excellent job in inspiring the investigation of situations where such feature-overlap costs occur. Yet, there are situations where they do not occur. For example, Craighero et al. (1999) observed faster discrimination of objects that correspond to a concurrently prepared grasping movement. Hommel and Schneider (in press) found facilitated detection of a visual target by concurrent selection of a feature-overlapping manual response, even at points in time when the response was ready to go, and thus according to TEC, its features were bound. Moreover, would TEC not predict that grasping an identified cup is a complicated task (cf. sect. 3.1.3)? The cup's identification comprises the binding of its features (including location) which are then *less* available for action control, thereby making a cup-oriented action difficult. To me, feature binding would be a more powerful explanatory principle if it was clear *a priori* when it will occur (and when it will not) – otherwise binding smells a bit like a post-hoc helper, recruited whenever feature-overlap costs call for explanation.

These inconsistencies might indicate that feature-overlap costs occur, at least occasionally, for reasons other than binding. They may originate from inhibition of feature codes (Caessens & Vandierendonck 2001), or because an already activated feature requires a larger activation increment to represent another event than at rest-activation level (Müsseler et al. 2001), or because feature-overlapping events are less discriminable as a result of their similarity (Duncan & Humphreys 1989). These considerations need not necessarily be correct but they may provide reasonable, albeit admittedly more traditional, theoretical alternatives.

***Movement-effect associations are usually context-sensitive.*** To be utilizable for goal-oriented action, movement-effect associations must often include the environmental conditions under which they are acquired (Hoffmann 1993). For example, a keypress on a PC-keyboard has dramatically different effects depending on the active program. A natural solution for this problem in TEC would be that movement-effect associations become bound with the stimulus conditions under which they are valid. Hommel et al. (2001) have shown that stimulus-response associations can indeed become "contextualized" in this manner. However, recent studies suggest that contextualization of response-effect associations is much harder to acquire, a fact that TEC should – and probably can – account for (Hoffmann & Sebald 2000; Stock & Hoffmann 2002).

Altogether, TEC has provided an extremely stimulating framework for the experimental study of goal-oriented action, and thus represents – even with some limitations in its present state – a major step towards uncovering the psychological reality of the hyphen in ideo-motor action.

ACKNOWLEDGMENTS

Funding of this work was provided by the German Research Foundation (Grant HO 1301/6-1).

## The CHREST model of active perception and its role in problem solving

Peter C. R. Lane, Peter C-H. Cheng, and Fernand Gobet

*ESRC Centre for Research in Development, Instruction and Training, School of Psychology, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom.*

**{pcl; peter.cheng; frg}@psychology.nottingham.ac.uk**
**http://www.psychology.nottingham.ac.uk/staff/{Peter.Lane; Peter.Cheng; Fernand.Gobet}**

**Abstract:** We discuss the relation of the Theory of Event Coding (TEC) to a computational model of expert perception, CHREST, based on the chunking theory. TEC's status as a verbal theory leaves several questions unanswerable, such as the precise nature of internal representations used, or the degree of learning required to obtain a particular level of competence: CHREST may help answer such questions.

The target article presents a unifying framework for perception and action, assuming their representation within a common medium. We discuss the relation of the Theory of Event Coding (TEC) to a tradition of computational modelling based on the chunking theory, which also addresses many of TEC's theoretical concerns; in particular, the notion of active perception guided by action sequences. The basic principles of TEC include:

1. Common coding of perceptual content and action goals.
2. Feature-based coding of perceived and produced events.
3. Distal coding of event features.

Principles (1) and (2) argue that action goals should be represented as composite action-events, just as perceptual objects are composite perceptual-features, and that integration is required to make perception an active element in action planning. Principle (3) implies that the level of representation is abstracted from that



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

of concrete actions or stimuli, facilitating transfer between perception and action.

The chunking theory (Chase & Simon 1973; Gobet et al. 2001) includes many of these elements and, further, embodies them within an established chain of computational models. The basic elements of the chunking theory are that collections of primitive features or actions frequently seen or experienced together are learned as an independently referenced chunk within long-term memory (LTM). Several chunks may be retrieved and referenced within short-term memory (STM), and their co-occurrence used to construct multiple internal references. One such internal reference is the association of perceived chunks with action chunks (Chase & Simon 1973; Gobet & Jansen 1994). These ideas have been extended into a computational model of diagrammatic reasoning, CHREST+ (Lane et al. 2000; 2001). One consequence of assuming that chunks underpin human memory is that output actions are driven by the perception of chunks in the stimulus; this leads to observable effects in the relative timing of the action events.

The CHREST (Chunk Hierarchy and Retrieval Structures) computational model of expert perception provides a unified architecture for perception, retrieval, and action (e.g., see description in De Groot & Gobet 1996; Gobet & Simon 2000). The model employs a discrimination network LTM, a fixed-capacity STM, and input/output devices, with timing parameters for all cognitive processes. Perception in CHREST is a process involving multiple interactions between the constituent elements of its architecture. The process may be summarised as follows: An initial fixation is made of the target stimulus and the features within the eye's field-of-view retrieved. These features are then sorted through the discrimination network so as to retrieve a *chunk* from LTM. This chunk is then placed within STM. The model next attempts to locate further information relevant to the already identified chunk; it does this by using any expected elements or associated links of the retrieved chunk to guide further eye fixations. Information from subsequent fixations is combined with earlier chunks, within STM, building up an internal image of the external world. If the LTM suggestions for further fixations are not of use, then the eye resorts to bottom-up heuristics, relying on salience, proximal objects, or default movements to seek out useful further features.

How is this process related to problem solving? As with the earlier models of chess playing, chunks within LTM become associated with relevant action sequences. One application of the CHREST model is to problem solving with diagrammatic representations; for example, solving electric-circuit problems with AVOW diagrams (details may be found in Cheng 1996; 1998); this version of CHREST is known as CHREST+). As CHREST+ acquires information about the two forms of stimuli, it constructs *equivalence links* between related chunks stored in LTM. During perception of a novel problem, CHREST+ identifies known chunks within the problem, and uses these to index chunks in the solution representation. These solution-chunks act as plans, forming the basis from which CHREST+ constructs an action-sequence composed of lines to draw. The level at which the CHREST+ model is operating is very much in concord with that assumed by TEC: the input is composed of features, representing the tail-end of primitive perceptual processing, and the output is a plan, representing the initiation of output processing. Both the input and output chunks are constructed from compositions of features, and both interact with the STM in a common representational format.

Although internal representations are difficult to investigate, a number of testable predictions may be derived from the chunking theory, relating to the relative timing of action sequences (Chase & Simon 1973; Cowan 2001; Gobet & Simon 2000). An example of the process is found when drawing complex shapes, where the latencies between drawing actions are partly governed by planning activities, and these planning activities are mediated by the state of the drawer's memory. A typical series of latencies includes a limited number of isolated local maxima, corresponding to longer periods of reflection; although participants are often not aware of their presence, these maxima are evident when detailed timing information is gathered. Such patterns have been shown to correspond well with a theory that chunks underlie the planning and memory processes (Cheng et al. 2001), and have also been modelled using CHREST+ (Lane et al. 2000; 2001).

Without detailed modelling of the lower-level empirical evidence presented in the target article, it is not fair to claim that computational models such as CHREST embody all of TEC's central ideas. However, what is interesting is that the aims of EPAM/CHREST, to capture the relatively high-level processes of expert memory in symbolic domains, have led to a model of active perception with a similar style to that proposed by TEC. CHREST may also be used to formalise some of the assumptions of TEC and turn them into empirical predictions. For instance, some of the harder areas of TEC to formalise are the form of the internal representation, or the amount of exposure to a specific domain required to learn a particular association between perceptual and action events. CHREST itself, with detailed timing parameters applicable across many domains, can be used to investigate such questions, using the model to derive strong predictions for the temporal information separating the output actions.

## Event coding, executive control, and task-switching

Nachshon Meiran

*Department of Behavioral Sciences and Zlotowski Center for Neuroscience, Ben-Gurion University of the Negev, Beer-Sheva, Israel, 84105.*
**nmeiran@bgumail.bgu.ac.il** **http://www.bgu.ac.il/beh/nachshon.html**

**Abstract:** Like the Theory of Event Coding (TEC), theories of executive functions depict cognition as a flexible and goal-directed system rather than a reflex-like one. Research on task-switching, a dominant paradigm in executive control, has revealed complex and some apparently counter-intuitive results. Many of these are readily explained by assuming, like TEC, that cognitive control is based on selecting information from commensurate representations of stimuli and actions.

There is a great deal of commonality between the Theory of Event Coding (TEC) and the study of executive control functions in general. First, both of these research trends are moving away from the reflex metaphor, which has dominated cognitive research for some time (e.g., behaviorism and additive factors). The reflex metaphor suggests that cognitive processing is similar to simple reflexes in that there is a linear progression from input (perception) to output (response). TEC moves away from the reflex metaphor in that it points out the many interfaces between perception and action. Theories of executive functions also move away from the reflex metaphor in that they perceive the cognitive system as an abstract general purpose structure that is being modulated according to changing task demands (e.g., Logan & Gordon 2001).

Second, TEC helps in understanding specific cognitive control functions. One example comes from the area of task-switching (e.g., Allport et al. 1994; Monsell & Driver 2000; Rogers & Monsell 1995). Typically, task-switching experiments involve stimuli which are equally relevant to all the tasks. For example, Fagot (1994) used colored letters and the two tasks were letter classification and color classification. In order to ensure that task-switching was performed at a central level rather than at a more peripheral effector level, the same set of response keys is often used in all tasks. For example, a given key might be used to indicate the identity of a given letter (e.g., "A") in the context of the LETTER task, whereas in the COLOR task the same key indicates RED. The critical variable in the task-switching paradigm involves comparing task-switched trials to task-repetition trials. Two additional variables are task congruency (whether the same physical response is indicated as the correct response in all tasks) and task repetition. The effects of advanced task preparation are also being studied.



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

The four variables just described produce a complex pattern of interactions in reaction time (RT). First, there are the main effects for task-switching (switch RT > repeat RT), congruency (incongruent RT > congruent RT), and task preparation (unprepared RT > prepared RT). In addition, there are three replicable 2-way interactions: The task-switching cost becomes smaller but is not eliminated as a result of task preparation (switch × preparation), congruency effects are larger in switch trials than in repeat trials (switch × congruency), and the usual response repetition effect (response-repeat RT < response-switch RT) is found only among task-repetition trials but is either not found or even sometimes reversed in task-switch trials (switch × response-repetition).

Perhaps the most counterintuitive effect found is the complete absence in most experiments of a triple interaction between switching, task-preparation, and congruency. The reason why this result is counterintuitive is the usual assumption made by many authors, that task preparation amounts to the replacement of one Stimulus-Response (S-R) mapping rule by another. However, this very reasonable assumption is apparently wrong since it predicts greater interference from the wrong mapping in unprepared trials, which would result in the above mentioned triple interaction. Thus, the effect of task preparation seems to involve something else – not the change of an S-R mapping rule.

The mathematical model developed by our group (Meiran 2000a; Meiran et al. 2000) explains all of these interactions in RT quite successfully, and some of its critical predictions were supported empirically (Meiran 2000b). Its assumptions are very much in the spirit of TEC. For example, it assumes a close interaction between abstract representations of stimuli and abstract representation of responses in the course of response selection. This interaction is so close that response selection is based on determining how "similar" the target stimulus is to each of the responses. Task control in the model is based on selective attention but, in the spirit of TEC, it assumes that selection operates both on the stimulus side and on the representation of the responses. For example, a LARGE SQUARE target stimulus can be represented, mentally, so that SQUARE is emphasized while LARGE is de-emphasized. Similarly, a key-press indicating both LARGE and SQUARE can be represented, mentally, as indicating SQUARE more than it is indicating LARGE.

The final aspect of the model constitutes a slight deviation from TEC, at least as TEC is presently formulated. However, I assume TEC can accommodate this deviation without changing any of its more substantial aspects. Specifically, in the model, the cognitive process responsible for selecting which information will represent a given physical response (termed "response task-set") is based on linking the physical response (e.g., pressing the right key) with the nominal response it is meant to indicate (e.g., RED). This process is similar to the formation of an event code in TEC, except for the fact that the "event code" in our model does not contain a perceived action outcome (a perceived key-press) but, instead, represents an intended nominal response (e.g., RED). The rationale is that pressing a given key serves as a substitute to saying, for example, "this letter is RED."

One potential objection to the above argument is that all our published modeling work involves switching two tasks, UP-DOWN and RIGHT-LEFT. Importantly, the responses (e.g., the upper-right key), like the target stimuli (e.g., the upper-left corner), contained both vertical and horizontal information. One could therefore assume that directing selective attention to a given dimension also affected how the position of the key-press was perceived and that the event code integrates the physical response with its perceived position, as it has been biased by selective attention. However, this interpretation is unlikely to be correct since our model also fits the results using paradigms in which switching was between a CIRCLE-SQUARE task and a SMALL-LARGE task ($R^2$=.98; Meiran et al., in preparation), and attending to a given target dimension was unlikely to affect how the key-press was perceived. Given the similar results in the two versions of the paradigm, a more parsimonious assumption is that the event code integrates the key-press with its associated nominal response rather than with its perceived outcome.

## A theory of representation to complement TEC

Ruth Garrett Millikan

*Philosophy Department U-54, University of Connecticut, Storrs, CT 06269.*
**ruth.millikan@uconn.edu**
**http://vm.uconn.edu/~wwwphil/faclist2.html**

**Abstract:** The target article can be strengthened by supplementing it with a better theory of mental representation. Given such a theory, there is reason to suppose that, first, even the most primitive representations are mostly of distal affairs; second, the most primitive representations also turn out to be directed two ways at once, both stating facts and directing action.

**Hommel et al.**'s Theory of Event Coding (TEC) in the target article can be strengthened by supplementing it with a better theory of mental representation. No explanation of the term "representation" is offered by the authors, but the quotations from Dretske (1981), the willingness sometimes to interchange the words "representation" and "information," coupled with statements such as "[t]hese codes are activated by, and thus represent, external stimulus information," strongly suggest some fairly simple causal or informational theory of representation. But Dretske (1986; 1988; 1995), along with many others, now recognizes that there can be no simple equation between mental representation and the occurrence in the brain of "natural information." We must add, at least, that a mental representation has representing as its function, using a sense of "function" in which functions can sometimes fail to be performed. Misrepresentation must not be ruled out by definition. This is usually taken to require that mental representations have been "designed," either by learning (Dretske 1988) or by natural selection or both, to represent what they represent, though, of course, what has been designed to do X may sometimes fail to do X. Better, the mechanisms that produce mental representations have been designed, by learning or natural selection, to produce and/or to learn to produce, representations that vary systematically so as to parallel variations in what needs to be represented (Millikan 1984; 1996; forthcoming). At first glance this may seem just a trivial addition to a causal or informational account of representation, but it is not. If the mechanisms that produce representations have been selected for by learning or by natural selection, this must have been owing to some systematic effects that the representations had, that is, to the uses to which these representations were put. Then it will be these uses that define the semantic contents of the representations, not their causes, and not whatever natural information they may or may not carry.

Now **Hommel et al.** "believe that abstract, distal representation has evolved as a solution to the problem of developing a representational scheme for the planning of goal-directed action" (sect. 5.2). But, from the above it follows that mental representations, even of the most primitive kinds, are most likely to represent distally. It is, for example, variations in the configurations of physical objects in one's immediate environment that have to be taken into account in order to appropriately guide most immediate physical activity, not the shifting quantities, qualities, and patterns of light striking the retina. Putting natural information to one side, what is mentally represented is only what natural selection has found it was useful to have represented. What a mental representation represents depends on how it is interpreted, rightly or wrongly, by the system that uses it. Consider the desert tortoise that automatically moves toward anything green. This disposition, hence the underlying neurology, has been selected for because it often effects movement toward vegetation. What the turtle per-


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

ceives, whether correctly or illusorily, is vegetation, when you might just see green. I am not talking about phenomenology, of course. I am talking about the semantic content of the turtle's mental representation. For the desert turtle, perceived vegetation may be very distal indeed, perhaps hundreds of yards away. Nor need we suppose that the turtle engages in "the planning of goal-directed action." Representation of distal affairs as needed to guide the organism's activity is the norm for nearly all perception, not a special quirk of human perception.

**Hommel et al.** tell us that "the main function of [the representation of an action] is to constrain the motor system to produce the action-contingent perceptual effects the action aims at" (sect. 4.2). It is hard to interpret this claim as consistent with a causal or informational theory of mental representation. I have argued, however, that there are two opposing directions in which mental representations may operate. They may be designed to represent facts they are not responsible for creating but which need to be taken into account to guide (physical or mental) activity. Or they may be designed, like blueprints, to help create the affairs that they represent. All of the most primitive representations, however, face both these directions at once. I call them "pushmi-pullyu" representations (Millikan 1996) after Hugh Lofting's charming creatures having that name. My favorite example is the bee dance, which tells the bees where the nectar is (facts) but tells them where to go as well (directions). This is also the realm of Gibsonian affordances, where certain invariants detectable in the energy surrounding the organism, or the "resonances" they set up in the organism, represent facts concerning certain features of the distal environment, but also direct how to respond to these features. Similarly, consider any explicit human intention. What you truly intend to do (not merely to try to do) must be something you believe you will do, but at the same time the intention's function is to guide you to get it done. If you forget your intention, for example, then you will not accomplish your intention. The very same mental entity or state (token) represents a fact about your future around which to adjust your other plans, as well as represents a plan which helps in creating that very fact. Ordinary language also contains pushmi-pullyu representations. My favorite is "No, Johnny, we don't eat peas with our fingers."

If **Hommel et al.** adopt a view of representation such that semantic value derives from the function rather than from the cause of the representation, they will see how plausible it is that not only do "perceived events and planned actions share a common representational domain," but that in many cases the same complete representation token can have two functions at once, being both a fact-presenter and an action-director. Indeed, the problematic issue now becomes, not "how did perceived events and planned actions come to share a common representational domain?" but "how did it happen that primitive representations, which always have both a fact representing side and an action directing side, come apart in the human mind?" The puzzle is how we eventually became able to entertain beliefs for which we may find no practical uses and also desires that, given our beliefs, we know cannot be fulfilled.

## Anomalous processing in schizophrenia suggests adaptive event-action coding requires multiple executive brain mechanisms

Robert D. Oades and Katja Kreul

*Clinic for Child and Adolescent Psychiatry, University of Essen, Essen 45147, Germany.* **oades@uni-essen.de**
**http://www.biopsychology.uni-essen.de**

**Abstract:** The integration of perceived events with appropriate action usually requires more flexibility to result in adaptive responses than **Hommel et al.** report in their selective review. The need for hierarchies of function that can intervene and the existence of diverse mediating brain mechanisms can be illustrated by the non-adaptive expression in psychiatric illness of negative priming, blocking, and affective responses.

Representatives of each arm of the firm gather around the conference table: there are those who interact with the outside world, the product makers and the salesmen, as well as those who organise their work, the production managers and market researchers. Not missing, of course are those with an overview over large parts of the firm's activities, namely, various directors ("executives"). We generalise about the function of each of these representatives, although in fact some have more abilities than others, and most are young enough to acquire more skills.

**Hommel et al.** are critical of Sanders (1983, p. 7) chart which stretches from stimulus processing to motor adjustment (interactions with the outside world), and flows over feature extraction and response choice in the hierarchy of synthesis and rationalisation (organisation and executive management). But they seem never to be quite able to break free from this historical precedent. The metaphor of the firm still seems to apply.

It is important for **Hommel et al.**, for example, that the binding of features and knowledge via selective attention to make an object file (Kahneman et al. 1992, p. 29), is the sort of psychological process operating in the perceptuo-motor interface which renders a distinction between perception and action functionally irrelevant. (Though perhaps, an anatomical distinction might still be tenable.) Does it matter whether our metaphorical conference actually occurs around a table at one locus or function (thanks to a software-based network) between topographically distributed participants? We think it does not usually matter. Indeed, in a similar vein, the authors cite indications that gamma oscillations seem to reflect binding functions even between disparate sensory and motor areas (Roelfsema et al. 1997). We made no mention of a specific product for the firm: for some products there may be roles or functions that are best united in one capacity, for instance, in product management and market research; just as, neurophysiologically speaking, they could be united in the functions of a "mirror neurone" receptive to input about the sorts of motor action in whose initiation it also participates (Gallese et al. 1996, p. 12). But are these not exceptions to the usual organising principles?

There may be a sense in which the Theory of Event Coding (TEC) can be tested, an aspect about which **Hommel et al.** were pessimistic (Introduction), if one does not neglect the anatomy and the mechanisms (cf. numerous versions of short/long feedback loops and reafference: e.g., the monitoring of action in Fig. 5.1 of Frith 1992). Hommel et al. state, "Not only should object and event perception be mediated by object-file types of structures, but action planning should be mediated by the very same entities as well . . . action files"; (sect. 4.2). Problems that concern some psychiatric patients (schizophrenia, obsessive compulsive disorder, OCD) can involve situations crucial to stages in the development of the TEC argument. The following three examples on the influence of irrelevant stimulation in mental illness, in priming the differentiation of central executive mechanisms rather than their unity, as required by TEC.

First, negative priming (NP) is said to provide pivotal evidence for the paradigmatic construction of an object file from subordinate perceptual features applicable to the integration of attention and intention. An interruption of the normal interference of the prime (e.g., in OCD; Enright & Beech 1993) would appear well suited to explain checking compulsions. But in fact checking does not relate to impaired inhibition in NP, but rather to facilitation on positive priming trials (Ferraro et al. 1996; Hartston & Swerdlow 1999). It seems counter-intuitive to expect that a correctly assembled object-action file should enhance the pathology without recourse to re-modelling several components of the process and their interactions. Interestingly, in the normal population impaired NP may relate more to schizotypal features (Sturgill & Ferraro 1997).



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

NP is impaired in acutely ill patients with positive symptoms of schizophrenia (Peters et al. 2000; Williams 1996), but is otherwise present in the disorganised and those with negative symptoms. This brings us to the second point. **Hommel et al.** (sect. 2.3.3.2.2) write that although actions performed in sympathy can involve involuntary automatic processes, they are often intentional responses to a situation that reflects object-action plan integration. We would predict that if NP (and the all important binding mechanisms, Haig et al. 2000) are impaired in positive schizophrenia, then inappropriate affect and sympathy should be a feature of these patients, which is not the usual finding. In fact, it is a cardinal symptom of the "disorganised" type of schizophrenia that may, but often does not involve the primary symptoms of positive schizophrenia (Straube & Oades 1992).

Our third point concerns **Hommel et al.**'s report that the direction taken by a masked arrow was less accurately reported when it concurred with left or right key responses than in the mismatch condition. Rather than representing an example of the influence of action plans on perception, we believe this may be better described as an example of conditioned blocking. That is to say, training with (the conditioning of) one response makes it less likely that the *wholly redundant* information associated with another stimulus is acquired. (The arrow is superfluous, the direction it points toward is not or only poorly registered for recall.) This is an entirely attention-related phenomenon seen in many forms of learning. It is also important to note that the details of the brain mechanisms involved in blocking differ from those associated with NP. This is reflected by reports that patients with negative or nonparanoid forms of schizophrenia are impaired on blocking (Bender et al. 2001; Oades et al 1996), not those with the positive form of the illness as in NP.

We conclude that it may be the exception rather than the rule that object files representing the binding of attentionally derived information achieve the functional identity of action files. Our argument is based on the different problems encountered by patients with one or another form of mental illness, and this implies the differential ("executive") intervention of the brain mechanisms involved in each of these phenomena. The impairments resulting from brain damage or psychopathology point to the diversity of these components and the need for anatomically diverse executive mechanisms, often based in the frontal lobes, to provide the link(s) to appropriate action plans.

## The role of feedforward control in motor planning

Marta Olivetti Belardinelli[a] and Demis Basso[b]

[a]*ECONA (Interuniversity Center for Research on Cognitive Processing in Natural and Artificial Systems) and Department of Psychology, University of Rome "La Sapienza," I-00185 Rome, Italy;* [b]*Department of General Psychology, University of Padua, I-35131, Padua, Italy.*
**belarditecona@iol.it** **demisbasso@supereva.it**

**Abstract:** In dynamical systems models feedforward is needed to guide planning and to process unknown and unpredictable events. Feedforward could help Theory of Event Coding (TEC) integrate control processes and could model human performance in action planning in a more flexible and powerful way.

Our commentary focuses on the relationship between control process and event coding mechanisma in motor planning, looking at this through the frame of a dymanic model of reality construction that considers subject and object as subsystems interacting within a single organism-environment psychological system (Olivetti Belardinelli 1974/1986). In this, all cognitive processes are considered as re-equilibration processes in which equilibrium retrieval is pursued via different modalities, referring either to different sensory channels, or to the influence of the subject's stable internal states. The latter aspect tunes the processing of discrepancy information in input and the choice of re-equilibration strategies to be executed in output. The possibility of spontaneous modifications of information processing modalities as complexity increases is guaranteed by feedforward (Stagner 1977), described as the proactive effect of the discrepancy information (Olivetti Belardinelli 1981–1982).

We consider the construction of reality as a starting point and all cognitive processing as a unique internal activity, strictly dependent upon the intermodal configuration of input information (Ferlazzo et al. 2001; Olivetti Belardinelli et al. 2001).

The Theory of Event Coding (TEC) is based on the hypothesis of an amodal working space, common to both representation of data and formulation of to-be-produced actions. A first problem concerns the level of this common working space, shared by early cognitive antecedents of action and late cognitive products of perceptual processing. The distinction between "late perception" and "early action" implies the existence of: (1) an "early perception" (difficult to allocate outside a proximal location); (2) a common cognitive space integrating the early and late perception.

Moreover, TEC gives too little consideration to control processes (Atkinson & Shiffrin 1968; Shallice 1982) that must necessarily be separated from the planning process. As it is not automatic that the processing of stimuli features will produce the performance of an action-response, the general control process must occur at an upper level with respect to planning.

A mechanism univocally linking a configuration of stimuli features to a response is only a partial description of human behavior in planning. As Norman and Shallice (1986) point out, a pragmatic theory should explain why a subject is able to choose another solution even if one is readily available (but does not satisfy him), and what happens when that subject determines his choice on factors external to the situation in input.

Furthermore, a theory should explain possible variations occurring during the execution of the plan. For example, in the paradigm of the Traveling Salesman Problem (a task in which visuospatial planning is involved, Cadwallader 1975; Hayes-Roth & Hayes-Roth 1979), the goal is to visit all the given locations through the shortest path. Subjects have to determine which order among $n$ locations optimizes the total travel distance. However, what happens in real life if we discover along the path a shop displaying a wonderful jacket but we don't have enough money to pay for it? We have to return home or go to a bank, and then revisit the shop. In other words, the plan must be rearranged by including this new information (Basso & Bisiacchi 1999). We cannot avoid enclosing a mechanism devoted to modifying plans in the moment of their implementation, given the likelihood that unpredictable events may occur.

According to TEC, the verification of action plans, the control of efficacy and efficiency, and the economy of cognitive load, are all determined by the plan creation. It seems an incomplete view to consider all these as simple parameters in the course of action planning. Although pre-existing action schemas may be stored in long term memory and directly activated by their self-organization, this is not enough to explain a re-planning consequent to unexpected events.

The need to verify the efficacy and efficiency of the just-conceived plans before acting seems to require a feedforward process that guides actions sequentially from perception to action by representing the final desired state (Fabre-Thorpe et al. 2001; Movellan & McClelland 2001). Whether stimuli are recognized or not, goal-directed behavior starts in order to retrieve the system's equilibrium by performing purposeful actions (Domjan et al. 2000; Freeman & Oster 1998; Thayer & Lane 2000). Vice-versa, in the case of unknown stimuli, the feedforward mechanism works by means of refinements determined by successive comparisons between actual and target product, in this way generating a learning process (Kringelback & Engell 1997).

To restore both new and known input discrepancy, an on-line control of execution is needed in the course of action (Della Malva


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

et al. 1993): moreover, in the case of external, unattended events, the feedforward allows the existing plan to be rearranged, adapting it to the events as soon as they happen.

A convincing attempt to model dynamic changes in "perceptual learning" can be found in Van Leeuwen and Raffone (2001). Their model enables flexible representations that capture the functional interplay of perceptual and memory representations, or, in other words, of bottom-up and top-down processing, thanks to a combination of the classic Hebbian learning and intrinsic chaotic dynamics. We maintain, therefore, that this approach could lead to the modeling of planning modifications in new and unknown situations (see, e.g., Lamme 2000; Payne et al. 1996).

In the Epilogue, section 5, the authors conclude that the construction of reality is an implication of the invention of identical codes for perceived and to-be-produced events. Our starting point (the construction of reality) is their conclusive construct. It is therefore possible that our two approaches are complementary, and analyzing planning processes from both perspectives may be needed in order to obtain a complete explanation of human motor in complex situations.

## Theory of event coding: Interesting, but underspecified

Chris Oriet, Biljana Stevanovski, and Pierre Jolicoeur

*Department of Psychology, University of Waterloo, Waterloo, Ontario N2L Canada.* **coriet@watarts.uwaterloo.ca**
**b2stevan@watarts.uwaterloo.ca** **pjolicoe@cgl.uwaterloo.ca**
**www.cgl.uwaterloo.ca/~pjolicoe**

**Abstract:** The Theory of Event Coding (TEC) is a new framework for understanding interactions between perception and action. We are concerned that the theory is underspecified, showing that it can easily be used to make exactly opposite predictions. Precise specification of the time course of activation and binding is needed to make the theory useful for understanding the perception-action interface.

***Common-coding versus classical theories of perception-action.*** One of the key principles that underlies the Theory of Event Coding (TEC) is that perception and action use a common coding format, thereby rejecting a more classical separate coding for perception and action. We are uncertain whether a common coding model can be distinguished from a classical model in which interactions between perception and action codes are postulated (Fig. 1). While the classical approach tends to view the relationship between perception and action as unidirectional, the additional assumption of reciprocal connections between these systems (upper panel) is sufficient to account for observed behavioural data without invoking the need for a single perception-action code (lower panel). **Hommel et al.** suggest that a common coding format is supported by the report of "mirror neurons" (e.g., Di Pellegrino et al. 1992) and "visual-and-motor neurons" (e.g., Sakata et al. 1995). Although it may be possible to distinguish between the two models on a *physiological* basis, it may be impossible, or at least very difficult, to do so based on *behavioural* data.

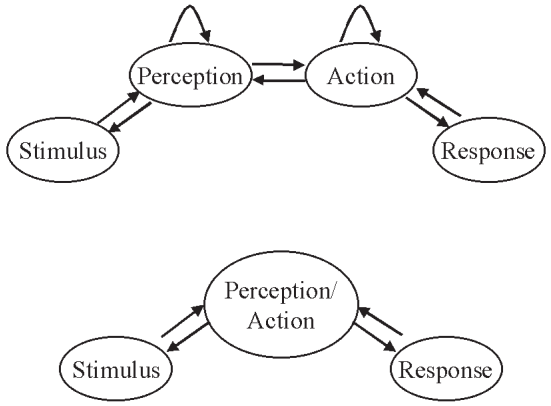



Figure 1 (Oriet et al.). Conceptualisations of the perception-action interface.

Using the Müsseler and Hommel (1997) paradigm, we have obtained results that conform to the TEC notion that the features of event files are coded abstractly. When the symbol ">" was interpreted as a right-pointing arrow, identification of that symbol was worse when it was presented during a right response than during a left response. When the same symbol was instead interpreted as a left-shining headlight, identification of that symbol was worse when it was presented during a *left* response than during a right response. Thus, as suggested by the TEC, what was required for coding the perceptual event was an abstract feature determined by the perceiver's intention to see either a left or right-featured target (Stevanovski et al. 2001a; 2001b).

***Time-course of action plans.*** Another finding from our lab, however, does not fit so easily within the current version of TEC. In sections 4.2.1 and 4.3.2, **Hommel et al.** refer to results published by Stoet and Hommel (1999) and Wühr and Müsseler (2000) where it is asserted that event codes are unbound immediately after they are used to make a response. In these studies, featural overlap with a recently executed response facilitated action and had no effect on perception. Because impairment was not observed, the authors took this as evidence that feature binding in the action plan had disintegrated so that no interference occurred. The authors suggest that, unlike feature binding in perception, the binding of features in planning an action dissolves "immediately after use." We have shown repeatedly that significant impairment can be observed for stimuli presented up to 750 msec after execution of an action (Stevanovski et al. 2001a; 2001b). These results cannot be reconciled with the notion that action files quickly disintegrate once the action is executed. The principles of TEC force us to conclude that the intention to perform an action was maintained (therefore the action file remained intact) beyond the point at which the action was actually executed. However, this conclusion, and the data upon which it is based, clash with previous results and theory. Our tentative conclusion is that the conditions that lead to facilitation versus inhibition of performance are not as well understood as suggested in the present version of TEC.

***TEC: Underspecified?*** The TEC is an interesting new framework for understanding the perception-action interface and for accommodating interesting results. Its greatest strength, however, is also its greatest weakness: it is difficult to imagine a finding in the action-perception literature that the theory *couldn't* accommodate. One reason for this is that the theory is underspecified: when costs in action or perception are observed, they are interpreted as reflecting the binding of event codes; when benefits are observed, these same features are assumed to be unbound, but activated. Although the authors posit some hypothetical time-course constraints of feature activation and binding, there are too many degrees of freedom to pin the theory down. For a rather striking example of TEC's underspecification, consider the results of Schubö et al. (in press). In section 4.3.2, the target article presents additional evidence for the influence of response generation on stimulus encoding. The authors report that effects of assimilation (and not contrast) were found between features of an executed response and a to-be-encoded stimulus. Despite the obvious similarities between this paradigm and the arrow paradigm described above, the authors argue that TEC does *not* invite competition and conflict for feature codes, but rather "clearly invites assimilation of the two codes." According to the authors, assimilation should be expected because action plans should have disintegrated at the end of the trial leaving some activation remaining in the feature codes. Because this lingering activation coexists with



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

the memorised codes, an assimilation of the codes occurs and subsequent memory is distorted.

Schubö and her colleagues, in fact, did not find evidence for assimilation (Schubö et al., in press, Experiment 4). Rather, contrast effects were observed: when the movement's amplitude was large, the stimulus' amplitude was encoded as smaller; when the movement's amplitude was small, the stimulus' amplitude was encoded as larger. Analogous effects were also found for the velocity of the movement. Given that the Schubö et al. (in press) results were contrasting, rather than assimilating, we would expect that the *contrast effects* would be interpreted as reflecting the binding of features into an action plan, which prevented access to that feature for the encoding of the stimulus. As illustrated by the preceding example, TEC can just as easily account for one pattern of results as it can for the exact opposite pattern, which we find somewhat troubling.

TEC is a first step in the right direction in that it has stimulated interesting investigations on the interaction between perception and action. However, the theory will have much greater utility when it is more focussed and able to specify *a priori* the time-course of feature activation and binding, and the conditions that produce interference and facilitation.

## Perception, action, and motor control: Interaction does not necessarily imply common structures

L. Pisella,[a,b] A. Kritikos,[b] and Y. Rossetti[a]

[a]*Institut National de la Santé et de la Recherche Médicale, Unité 354, "Espace & Action," 69676 Bron, France;* [b]*School of Behavioural Science, Department of Psychology, University of Melbourne, Victoria 3010, Australia.*
**pisella@lyon151.inserm.fr** **a.kritikos@psych.unimelb.edu.au**
**rossetti@lyon151.inserm.fr**
**http://www.lyon151.inserm.fr/unites/534_rossetti.html**

**Abstract:** The Theory of Event Coding (TEC) provides a preliminary account of the interaction between perception and action, which is consistent with several recent findings in the area of motor control. Significant issues require integration and elaboration, however; particularly, distractor interference, automatic motor corrections, internal models of action, and neuroanatomical bases for the link between perception and action.

The spectrum of empirical evidence selected in support of the Theory of Event Coding (TEC) is not sufficiently broad to form the basis of a general theory of perception and action. First, the authors are mainly referring to key pressing and imitation. A theory of action should encompass the most common daily interactions with the environment: goal-directed actions and bimanual co-ordination. Second, the authors specify that TEC only deals with action planning, excluding "the late motor processes that subserve motor realisation." Does action planning only refer to goal selection? If so, it is not surprising that perception and action overlap. The actual challenge for TEC is to integrate the field of motor control. We argue below that there are close links between perception and action in motor execution as well as in motor planning, as evident in paradigms of distractor interference and target perturbation. The links may manifest as avoidance (distractor inhibition) or attraction (automatic correction) in the trajectory of the hand.

A major gap in TEC is a satisfying account of the nature and role of distal codes, whether in button presses (in reaction time paradigms) or in goal-directed actions (pointing and reach-to-grasp).

First, the hand starting position relative to environmental stimuli influences action planning (Rossetti et al. 1994) and execution (Tipper et al. 1998). How does TEC account for this in the context of action represented at the level of distal feature codes?

Second, how does TEC account for distractor interference? Particularly, are distractor codes suppressed to result in successful action? Tipper and colleagues (e.g., Tipper et al. 1998) emphasise that excitatory activation to the target is modulated by competing inhibitory activation towards the distractor, accounting for alterations in the curvature of trajectories of the hand to the target. The concept of excitatory activation to target features is consistent with TEC. The inhibitory effect of distractors, however, indicates a complex and crucial role for irrelevant stimuli not addressed in TEC. Furthermore, the notion that feature codes are "glued" together, and that the similarity between codes may facilitate or interfere with actions, needs to be placed in the context of influential work by Duncan and Humphreys (1989; 1992). Preliminary work (Kritikos, in preparation) indicates that this principle may also apply to reach-to-grasp movements: time to peak velocity is altered when the target and distractor have features in common (e.g., colour, shape), showing that interference depends on the degree of target and distractor similarity.

In the field of motor control, action is also dependent on both internal and external causes (sect. 2.2.3.1). We have studied the relationships between automatic (stimulus-driven) and intentional (cognitively-driven) motor reactions to target perturbations (Pisella et al. 2000). The results reveal two discrepancies with the bases of the TEC.

First, when the authors claim that perceptual and motor codes are the same, do they imply that these codes are based on the same processes and structures? If so, how can TEC account for the fact that the same perceptual event can give rise to several different actions (e.g., designating, grasping, squeezing)? Does the perception of a distal stimulus imply that one has already selected a given action to perform on or with this stimulus? It seems more coherent to conceive of several dissociated sensory and motor processing streams emerging from one stimulus, as a result of automatic (spontaneous) or intentional (task-dependent) links. Thus, the activation of one "feature-code" activates a large sensory and motor network. Indeed, activation of a hand automatic pilot in a location-stop condition (see Fig. 1) revealed such a spontaneous and irre-

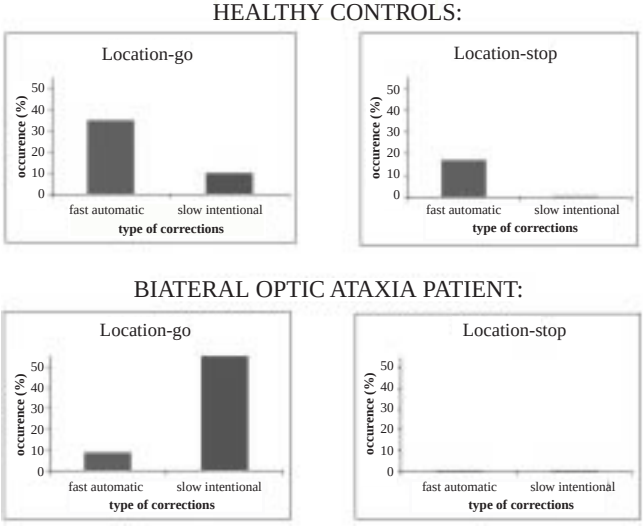



Figure 1 (Pisella et al.). Healthy controls and a patient with bilateral optic ataxia performed a pointing task where the target could change location at movement onset. In response to this target perturbation, the instruction was to correct (location-go condition) or to stop (location-stop condition) the ongoing movement. Fast automatic corrections were irrepressibly performed by healthy controls in the location-stop condition. They performed additional slow intentional corrective movements in the location-go condition. Bilateral lesion of the dorsal stream was associated with a specific impairment of the fast automatic corrections. (Adapted from Pisella et al. 2000)


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

pressible link between specific sensory and motor networks. Comparison of the temporal constraints of different stimulus-response combinations showed temporal additivity between sensory and motor modules (Pisella & Rossetti 2000), suggesting a dissociation between the sensory processes of the location change and the fast motor corrective processes of the hand automatic pilot.

Second, automatic corrective movements contradict **Hommel et al.**'s claim that stimulus presentation is necessary but insufficient for response to occur (sect. 2.2.3.1). Initiation of responses which are not entirely pre-programmed (i.e., different from key-pressing, as in Taylor & McCloskey 1990) can hardly be unintentional; whereas movement control can be driven unintentionally (Rossetti & Pisella 2001). We distinguish a system of automatic motor control (dorsal to premotor) from a system of conscious perception and intentional action (ventral to prefrontal). In the latter, perception and action directly share common processes and structures. In the former, we argue that the processes underlying automatic action are internal models, which represent an indirect convergence between perception and action, as developed below.

Theories of motor control postulate the existence of internal models subserving anticipatory processes used in everyday actions (Frith et al. 2000) and necessary to perform fast automatic corrections (Desmurget & Grafton 2000). In these models, the goal of action is not the perceived target location, but the final posture allowing this goal. Internal models therefore emphasise goal representation and possibly provide the background link between "distal stimuli" and "action parameters" alluded to in TEC. For they are intermediate between action and perception: they are built and continuously adapted through the perceptual effects of previous motor acts. We think that they represent the sensory and motor knowledge about spaces and objects, and are relevant not only for motor control but also for perception and cognition (Frith et al. 2000).

To sum up: The issue of an anatomical substrate for a common code for perception and action is poorly addressed by TEC. For example, the dominant theoretical framework of a ventral and dorsal stream distinction (Milner & Goodale 1995) remains an important background. It continues to evolve with accounts of temporal constraints and interconnections of the two streams. Recent interpretations challenge the initial dissociation between perception and action and emphasise integration between these processes (Milner & Dijkerman 2001; Rossetti & Pisella 2001; Rossetti et al. 2000). The parietal lobe and the cerebellum are two prominent examples of integration that are crucial in internal models (Desmurget et al. 2001; Frith et al. 2000; Imamizu et al. 2000); yet they have been largely ignored in the target article.

## TEC: Integrated view of perception and action or framework for response selection?

Robert W. Proctor and Kim-Phuong L. Vu

*Department of Psychological Sciences, Purdue University, West Lafayette, IN 47907-1364.* **proctor@psych.purdue.edu**
**http://www.psych.purdue.edu/~proctor** **kvu@psych.purdue.edu**
**http://www.psych.purdue.edu/~kvu**

**Abstract:** The Theory of Event Coding (TEC) presented in **Hommel et al.**'s target article provides a useful heuristic framework for stimulating research. Although the authors present TEC as providing a more integrated view of perception and action than classical information processing, TEC is restricted to the stage often called response selection and shares many features with existing theories.

**Hommel et al.** argue that cognitive psychologists have tended to neglect the interface between perception and action. They stress the need to develop an adequate account of that interface and propose a specific framework based on coding of stimulus and response features. Most of our research has been devoted to coding explanations of stimulus-response compatibility and related effects (e.g., Proctor & Reeve 1986; Vu & Proctor 2001). Consequently, there is no major disagreement between our views and theirs on these fundamental points, and our commentary focuses on the specific claims and arguments that Hommel et al. make.

**Hommel et al.** assert, "Though a large number of cognitive psychologists subscribe to the information-processing approach, only a few of them have taken interest in the total information stream between stimuli and responses" (sect. 2.1.1). They imply that their research is an exception, stating that the Theory of Event Coding (TEC) provides "a framework that would allow a fully integrated view on a substantial part of human-information processing" (sect. 3). Moreover, according to them, "In contrast to the classical information processing view, it [TEC] does not see perception and action planning as different, functionally separable stages but as intimately related, sometimes even indistinguishable processes producing a wide variety of interactions" (sect. 1, para. 4).

However, **Hommel et al.** restrict perception to the "'late' cognitive products of perceptual processing that stand for, or represent, certain features of actual events in the environment," and action to the "'early' cognitive antecedents of action that stand for, or represent, certain features of events to be generated in the environment" (sect. 1, para. 3). The exclusion of early perceptual processing and late action processing means that TEC does not address most of what would be of primary concern to researchers in perception and motor control. Instead, the focus is on that portion of information processing typically attributed to a central processing stage called response selection, response determination, response choice, or stimulus-response translation. We see little difference in the distinction between early perception, late-perception/early action, and late action and the classical distinction between "perception, translating perception into action, and the controlling of action" (Welford 1976, p. 3). Thus, Hommel et al. are no less guilty than anyone else of partitioning human information processing into functionally separable stages and restricting their investigations to a specific stage.

The nature of response selection has been investigated since the field's earliest days (see Proctor & Vu, in press a). Studies by Hick (1952), Hyman (1953), and Fitts and Seeger (1953), among others, are among the most widely cited classics in contemporary human performance, and although many cognitive psychologists may not be well versed in this area, there has been continual theoretical development since the 1950s. Therefore, the most appropriate way to view TEC is as a framework for understanding this central stage of information processing, and TEC should be evaluated relative to other frameworks and models of response selection.

The stress on cognitive coding is not unique to TEC. **Hommel et al.** claim that the classical studies examining the interaction between stimulus information and information in memory "focus on processes rather than on contents and structure" (sect. 2.1.1.2), but researchers have focused on the nature of cognitive codes since the rise of the information processing approach. For example, Chapter 3 of Posner's (1978) classic book *Chronometric Explorations of Mind* is titled "Coordination of Codes," and the index shows that codes are mentioned on 46 of the book's 241 pages. Furthermore, coding accounts are widely accepted by researchers investigating response selection (e.g., Umiltà & Nicoletti 1990). In fact, two of the most influential approaches to response-selection in the past 20 years, the salient features coding framework (Proctor & Reeve 1986) and the dimensional overlap model (Korblum et al. 1990), rely on feature codes for both stimuli and responses. Although these and other coding approaches do not emphasize distal reference, it is now well accepted that responses are coded in terms of the action goals. Hommel et al. are critical of the translation metaphor often used to characterize response selection because it "stresses the incommensurability between stimuli and response" (sect. 2.1.1.2), rather than common coding. However, the more fundamental idea behind referring to the stage as "transla-



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

tion" is that the stage "incorporates the traditionally recognized *choice of response,* but also emphasizes the complex mediating processes that may be involved" (Welford 1968, p.18).

The emphasis on event codes in a common coding format is probably the most distinctive feature of TEC. This view links the research on response selection with that on attention and emphasizes that attention involves more than just binding stimulus features. Most research on response selection has focused mainly on the direction of perception-to-action. One consequence of adopting the common coding view is that the research of Hommel and colleagues has placed more emphasis on the direction of action-to-perception. A second positive consequence of the common coding view has been for Hommel and colleagues to devote more research to dynamic events than is typical in the response-selection literature. Additionally, the event file concept seems to provide considerable insight into repetition effects. As **Hommel et al.** note, it predicts that effects of response repetition will interact with repetition or change of stimulus features. Consistent with this implication, Hommel (1998b) found that performance was worse when there was a partial repetition (one stimulus feature repeated and another changed) than when either all features repeated or all changed. We have found this pattern of results when tasks with two different stimulus-response mappings are mixed (Proctor & Vu, in press b).

**Hommel et al.** base TEC closely on the ideomotor view of action because of their focus on action goals and a common medium for perception and action. They indicate that TEC takes elements from two theories, ideomotor theory and motor theories of cognition, both of which propose that the motor system plays an important role in perception and cognition. They contrast the ideomotor view with the sensorimotor view, which they propose underlies the linear stage theory of human performance. Yet, although it is clear that a theory in terms of coding stimulus and response features is necessary, it is far less clear that an ideomotor view of action is to be preferred. Hommel et al. note, "Sensorimotor theories of action have so far been clearly more successful than ideomotor theories" (sect. 2.1.1). Although they explain away this fact as due to the difference in the explanatory strategies of the two views, it should give one pause about embracing a variation of ideomotor theory. Similarly, the point mentioned in passing by Hommel et al., that they deliberately excluded the motor theory of speech perception "partly because empirical support seems weaker than the elegant principle suggests" (sect. 2.1.4.1), should not be taken lightly since that theory is probably the most strongly advocated and widely investigated exemplar of a motor theory.

In the last paragraph before the Epilogue, **Hommel et al.** emphasize that the central strength of TEC is its heuristic value for stimulating research. This point is beyond question, given the extremely large number of important empirical studies on perception-action relations that the Max Plank group has generated. Due in large part to their efforts, research on response selection has seen tremendous growth over the past decade and the critical role of central response-selection processes in human performance has come to be more widely appreciated. We have little doubt that their research efforts within this framework will continue to be highly informative, regardless of the ultimate fate of TEC.

## The TEC as a theory of embodied cognition

Daniel C. Richardson and Michael J. Spivey

*Psychology Department, Cornell University, Ithaca, NY 14853.*
**dcr18@cornell.edu** **http://susan.psych.cornell.edu/home**
**spivey@cornell.edu**
**http://www.psych.cornell.edu/faculty/people/Spivey_Michael.htm**

**Abstract:** We argue that the strengths of the Theory of Event Coding (TEC) can usefully be applied to a wider scope of cognitive tasks, and tested by more diverse methodologies. When allied with a theory of conceptual representation such as Barsalou's (1999a) perceptual symbol systems, and extended to data from eye-movement studies, the TEC has the potential to address the larger goals of an *embodied* view of cognition.

The two main claims of the Theory of Event Coding (TEC), as we see it, are the nature of perception-action relations and the role of distal coding. The former claim is supported with data from experiments that require overt behaviours, such as button presses, in rapid response to reasonably simple stimuli. Yet if it is the case that event codes are an important level of representation, then we might expect them to contribute to more complex, conceptual representations as well. Also, if common coding is indeed a key principle of mental representation, then we should find evidence of motor systems participating in a wide range of cognitive tasks. Following these assumptions, we propose that the TEC should be allied to broader theories of cognition such as Barsalou's (1999a) Perceptual Symbols Systems, and present several examples of how the TEC can be extended into methodologies that use richer stimuli and more complex responses. We argue further that problems associated with **Hommel et al.**'s second claim, regarding distal coding, are alleviated by adopting an embodied view of the mind.

Barsalou's (1999a) Perceptual Symbol Systems theory proposes an account of conceptual representation as perceptual simulation, or combinations of "perceptual symbols." Although motor systems are proposed to contribute to perceptual simulations, most of the empirical evidence so far focuses on the modal, perceptual quality of conceptual representations. For example, Stanfield and Zwaan (2001) showed that if a narrative implicitly describes an object in a certain orientation (e.g., a nail hammered into the floor versus the wall), an image of that object presented later will be identified more quickly if it is in the same orientation. This suggests that subjects construct a mental model or "perceptual simulation" of the narrative that is specific to the level of object orientation.

Recent work in our laboratory uses the framework of the TEC to show that perceptual simulations of a spoken narrative may also include "action codes" for potential motor interactions. In Experiment 2 of Richardson et al. (2001), subjects heard a story which mentioned an object with a one-sided affordance (e.g., a milk jug). The orientation of the object was implied (e.g., "the milk jug points towards the egg cup"). When subjects made a later judgment about a different property of the object (e.g., "Is the jug in the center of the table?"), there were interactions between the left/right hand used to make the response, and the left/right affordance of the imagined object. Thus, the mental representation of the scene seems to evoke action codes for object affordances, which then interact with manual responses. Moreover, the time course of these interactions followed the pattern suggested by TEC (sect. 3.2.2) such that early responses show stimulus-response compatibility effects, and later responses show incompatibility effects (Stoet & Hommel 1999; Tucker & Ellis 1998).

However, the speed of a button press is only one, rather limited, way to investigate the interaction between motor and cognitive systems. **Hommel et al.** frequently allude to eye movements as a concrete example of action continuously guiding perception, yet they do not discuss the literature on eye movements and cognition (e.g., Ballard et al. 1997). Work in our laboratory has shown how oculomotor systems participate in the comprehension of spoken scene descriptions (Spivey & Geng, in press; Spivey et al. 2000), and the spatial indexing of linguistic information (Richardson & Spivey 2000). In the latter work, we presented adults with auditory semantic facts that co-occurred with visual events in four locations. When answering a question about one of the facts, subjects made saccades to the (now empty) region of space associated with that fact. Although location was task-irrelevant, adults tagged auditory information with a spatial index (cf. Pylyshyn 2001). In this way, even though they may be looking at a blank display, subjects' eye movements reveal a degree of oculomotor participation in what is essentially a distal encoding of linguistic information.

Placing the TEC within the context of the embodied cognition



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

framework not only widens the scope of this promising theory, it also alleviates the metatheoretical concerns that **Hommel et al.** raise in their Epilogue about the source of distal coding. From an embodied perspective, the reason we do not see spikes traveling along our nerves when we look at a flower is because those spikes *are* the seeing. An embodied mind is one in which the entire system, from sensors to effectors, participates in mental experience. There is no subset region that one can point at and call "the mind within the body." Thus, when the TEC is combined with an embodied view of cognition, the question of why coding is distal instead of proximal fades away.

In conclusion, we see an increasing body of research that reveals a very tight coupling between perception and action. Within the domain of reasonably simple, stimulus-response tasks, **Hommel et al.** replace the idea of a perception-action arc with the elegant notion of event codes. We argue that this common coding structure may also be employed in conceptual or linguistic representation. If one views cognition from an embodied or situated perspective (Barsalou 1999b; Clark 1997; Glenberg 1999; Lakoff 1999), then language comprehension is a matter of "preparation for situated action." In this framework, conceptual representations are contiguous with the representational forms of perception and action. Therefore, with a broader perspective on cognition and a wider set of empirical methodologies, the TEC provides the theoretical mechanics for investigating the embodied nature of cognition.

ACKNOWLEDGMENTS

Supported by a Sloan Foundation Fellowship in Neuroscience to MJS, and a Sage Fellowship to DCR.

## The event-code: Not the solution to a problem, but a problem to be solved

Michael J. Richardson and Claire F. Michaels

*Department of Psychology, University of Connecticut, Storrs, CT 06269.*
**michael.richardson@uconn.edu**
**http://www.ione.psy.uconn.edu/~cespaweb/**
**claire.michaels@uconn.edu**
**http://www.ione.psy.uconn.edu/~cespaweb/**

**Abstract:** We commend the argument that perception and action are tightly coupled. We claim that the argument is not new, that uniting stimulus and response codes is not a problem for a cognitive system, only for psychologists who assume them, and that the Theory of Event Coding (TEC)'s event-codes are arbitrary and ungrounded. Affordances and information offer the common basis for perception-action (and even for event-codes).

We commend the argument of **Hommel et al.** that perception and action are tightly coupled and should not be dealt with in isolation; it is a principle not often prescribed by advocates of an information processing approach. However, this argument is not new, and has been asserted, elaborated, and its implications tested by ecological psychologists for some time now (see Gibson 1966; 1979; Kugler & Turvey 1987; Michaels & Carello 1981; Shaw & Bransford 1977; Turvey 1977). Therefore, it is perplexing that Hommel et al. have essentially chosen to dismiss the ecological approach because of the Theory of Event Coding (TEC)'s "emphasis on representational issues – which are anathema to ecological approaches by definition" (Introduction). Hommel et al. attempt, thereby, to carve out their own world, only briefly describing theories (not just the ecological approach) outside the scope of that world, with no definite attempts at integration. The world they carve out is populated by mental representations underlying what they term late-perception/early action, which, they argue, must combine the sensory and the motor. We point out several problems with this position.

TEC appears to be founded on the notion that separate sensory codes and response codes exist or have existed: "we suppose that at some unknown point in the evolution of animal minds, it was highly advantageous to develop a new representational domain for matching perception and action planning – on top of the old system" (sect. 5, Epilogue). The notion of "the old system," however, reflects a deep confusion between the explanations and explananda of psychology; separate stimulus codes and response codes are concepts made by psychologists and not part of the landscapes of evolution or learning. Indeed, evolution and learning have their own problems to solve; let's not suppose that phylogeny and ontogeny recapitulate psychology.

A second perplexing issue concerns **Hommel et al.**'s addressing perception-action interdependence only in a sphere they term late perception and early action. Indeed, one wonders what "early" perception and "late" action are, not to mention what the "middle" constitutes. If perception and action are so tightly entwined, then why attempt to build a theory to consolidate this notion by parsing the circularity about which this coupling is formed? The implied succession from early perception to late perception to early action and so forth seems precisely the linear perception-to-action language that the authors criticize in section 2.1. In fact, we are fearful that distinctions such as "late" and "early" simply mirror the dualism they are arguing against.

In rejecting the traditional dualism of perception and action, **Hommel et al.** formulate a representational position – event codes, which even we prefer to "the old system." Unfortunately, they pursue another traditional dualism with a vengeance, that of animal and environment. The consequence of treating the person (or animal) independent from the environment is that the responsibility for forming event codes falls squarely on the perceiver; it is an entirely mental process, grounded neither in reality nor in information. There is no event ontology and no information specifying those events. Instead, the perceiver must bind together arbitrary coding sets (comprising elements such as green, red, big, left, right) according to seemingly old-fashioned associationistic processes. In short, **Hommel et al.**'s union of perception and action is created by the artificial binding of elements of an underconstrained (and largely unidentified) roster of capriciously selected and ungrounded symbols (see also Michaels & Stins 1997).

If one cannot live without representations, then at least let them represent affordances and originate in information about affordances. Rooting the measurement of animal-environment relations in natural law and constraints offers the sought-after common basis of perception and action. Examples include the work by Gibson on the optic array (see Gibson 1979), Turvey and colleagues on dynamic touch (see Amazeen & Turvey 1996; Turvey 1996), as well as the works mentioned by **Hommel et al.** on climbability (Warren 1984) and the sit-on-ability (Mark 1987). Consequently, the correct application of the concept of affordance, which characterizes the animal-environment fit and encapsulates perception and action, would seem better suited for their aims. Similarly, TEC would benefit from exploiting the concept of information for the perceiver/actor, especially noting its specification sense.

Our usage of the term "correct application" reflects our opinion that **Hommel et al.** sometimes misrepresent the way ecological psychologists define and employ the concepts of affordances and information. Affordances are environmental properties taken in reference to the action capabilities of the animal. They do not mediate perception and action, nor do they evoke actions (as seems to be implied by Hommel et al.'s statement, "action gets directly tuned to perception"); rather, affordances are possibilities for action and are embedded in a broader theory that formally addresses intention and attention (see Kugler & Turvey 1987; Shaw & Kinsella-Shaw 1988). The role of the ecological concept of information, especially in its specificational sense is also misconstrued. For example, when discussing the work of Runeson and Frykholm on the perception of human action (sect. 2.1.4.1, para. 3), Hommel et al. state that "people are . . . capable of inferring



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

the kinematics or even the dynamics of the movements." As argued by Runeson and Frykholm (1983), one does not "infer" the dynamics of a movement event; one detects information that specifies it.

We believe that if one accepts the mutuality of perception and action and, likewise, the mutuality of animal and environment, based on natural law, then abstract mediation and codes become unnecessary. However, if theorists insist that representations or "event codes" are needed to explain perception and action, they must also provide a law-based explanation of what the code constitutes and what brings it into existence. Until that basis is provided, such a theory is void of explanatory power – an issue that **Hommel et al.** clearly need to address.

## Computational motor planning and the theory of event coding

David A. Rosenbaum

*Department of Psychology, Pennsylvania State University, University Park, PA.* **dar12@psu.edu**

**Abstract:** Recent computational models of motor planning have relied heavily on anticipating the consequences of motor acts. Such anticipation is vital for dealing with the redundancy problem of motor control (i.e., the problem of selecting a particular motor solution when more than one is possible to achieve a goal). Computational approaches to motor planning support the Theory of Event Coding (TEC).

My aim in this commentary is to underscore the importance of anticipating the consequences of motor acts when they are planned. Recent computational models of motor planning have relied heavily on this notion. Covering them more fully than **Hommel et al.** did provides additional weight to their argument.

Consider the ostensibly simple task of bringing the fingertip to a point in extra-personal space. The point has three degrees of freedom – its horizontal ($x$), vertical ($y$), and in-out ($z$) position – but the fingertip, viewed as the end-effector of a multi-link limb segment chain, has many more degrees of freedom. The shoulder has 3 axes of rotation, the elbow has 2, the wrist has 2, and the finger joints have still more. If the point to be touched lies within reachable space rather than at its edge, then by definition there is more than one way to contact the point with the fingertip. The fact that a single posture is adopted raises the question of how that posture was chosen. Similarly, because the intermediate positions on the way to the final position could also be reached with any of a number of postures, the question generalizes to how the entire movement path was specified. Considering the forces applied by the fingertip as well as the positions it adopts, the problem extends from kinematics (the analysis of positions in time) to dynamics (the analysis of forces in time).

These problems have been approached by computational motor-control researchers in ways that can be placed in a $2 \times 2$ table (Table 1). The rows of the table differ with respect to mapping direction and the columns differ with respect to type of computation.

Table 1 (Rosenbaum). *Approaches to movement planning*

| | *Computation* | |
|---|---|---|
| *Mapping* | Table Look-Up | Model |
| Forward | | |
| Inverse | | |

Regarding mapping direction, forward mapping takes as input joint angles (in the case of kinematics) or muscle forces and muscle torques (in the case of dynamics), yielding as outputs spatial positions (in the case of kinematics) or spatially referenced forces and torques (in the case of dynamics). Inverse mapping works the other way around. Inputs are either desired spatial positions for which appropriate joint angles must be found, or they are spatially referenced forces and torques for which appropriate muscle forces and muscle torques must be determined. Regarding computations, the table look-up method relies on retrieval of previously stored instances, whereas reliance on models entails formulas or their implicit representations.

Each approach in Table 1 has advantages and disadvantages. The table look-up method, which uses retrieval of previously stored instances, can speed planning when familiar tasks are encountered, but may falter when new tasks are required – though interpolation can be used to deal with new task demands (Atkeson 1989). Using models rather than tables allows for generativity, and evidence has been obtained for model-based generalization (e.g., Conditt et al. 1997). On the other hand, model-based approaches have at least two limitations. First, no one, to my knowledge, has shown how model complexity relates to learnability. Second, advocates of model-based planning have not yet explained why planning speeds up with practice.

Inverse mapping has intuitive appeal as an approach to motor planning because one typically has an external goal and one must "work backward" to determine how the goal should be achieved. On the other hand, inverse mapping often runs into a one-to-many assignment problem. More than one solution is possible, making the problem ill-posed from a mathematical standpoint. The forward problem is never ill-posed because a unique output is always possible, given a fully specified input. However, forward mapping is a more round-about solution to the normal planning task than is inverse mapping because one typically does not have as one's primary aim the determination of a limb position, given a set of joint angles, or the determination of spatially referenced forces and torques, given a set of muscle forces and torques. Because forward mapping yields unique solutions, however, it has become the method of choice in a number of accounts of motor planning (e.g., Rosenbaum et al. 2001). With this method, soft constraints may be added to limit the search to joint angles or to muscle forces and muscle torques that satisfy primary spatial constraints (bring the end-effector to the correct spatial location where it can apply the correct force), as well as soft constraints (move there gracefully).

As seen above, none of the methods summarized in Table 1 is perfect. Different ones may be used in different tasks, by different effectors, and conceivably at different stages of practice. Determining which methods apply in which situations has not yet received much attention, although the possibility of combining the approaches has been recognized (Wolpert & Kawato 1998).

This point aside, what all the methods share is strong reliance on anticipation of the consequences of prospective motor acts. Even when the consequences are mundane, as in determining the location in external space where a point along the limb segment chain will be given a possible movement, anticipation is vital, owing to the redundancy of the motor system (i.e., its potential for achieving a single goal in multiple ways).

A simulation that illustrates the importance of anticipation is shown in Figure 1. This figure shows an artificial creature grasping an object after reaching around an obstacle. The simulation was generated by relying on forward mapping of postures onto spatial positions as well as model-based determination of spatial positions given candidate postures. In the simulation, candidate postures were evaluated with respect to a constraint hierarchy for which obstacle avoidance was most important, grasping of the target was second most important, and minimization of angular excursions by joints requiring large amounts of energy was least important.

The ability of the simulated creature to perform this complex action attests to the importance of anticipation in everyday action



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

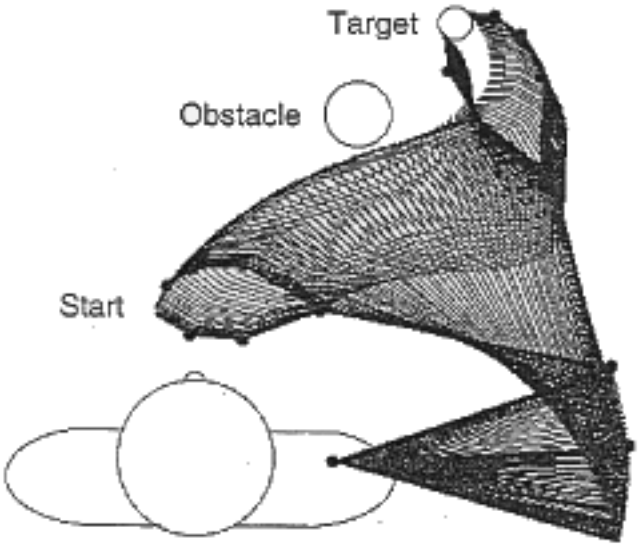



Figure 1 (Rosenbaum). Simulated grasp of an object after avoidance of an obstacle. From Rosenbaum et al. (2001).

planning. Research on computational motor control has borne out **Hommel et al.**'s claim at the end of their article that "abstract, distal representation has evolved as a solution to the problem of developing a representational scheme for the planning of goal-directed action."

## How specific and common is common coding?

Andries F. Sanders

*Department of Psychology, Vrije Universiteit at Amsterdam, Amsterdam, The Netherlands.* **af.sanders@wxs.nl**

**Abstract:** This commentary addresses three points. First, it is argued that the common coding principles, as developed in the target article, may supplement rather than replace stage views of human information processing. Second, the issue of the properties of an event code is briefly discussed. It is concluded that much remains to be specified so as to allow critical tests. Finally, the question of the limits of common coding is raised. It may be particularly relevant to direct perception and action coupling but less useful for the analysis of cognitive skills.

Perception and action-planning in a common representational medium has been a major *Leitmotiv* of the research carried out at the Experimental Psychology Department of the Munich Max Planck Institute for Psychology, ever since Wolfgang Prinz accepted the position of director there during the late 1980s. In principle, the research programme had already been defined in Prinz's (1983) book but a group at the Max Planck Institute provided the opportunity for a concerted research effort. The target article by **Hommel et al.** provides an important and well-written interim report of some of the more recent research. A considerable number of ingenious studies is reviewed, all inspired by the common coding hypothesis and all demonstrating its potential as an umbrella concept for perception and action planning. All in all, it is a clear demonstration of the excellent policy of concerted research programmes as favored by the Max Planck Gesellschaft.

Yet, critical issues remain. In this commentary I would like to briefly address three points. The first refers to the extent to which the authors, led by their enthusiasm, go too far in subsuming the wide variety of human information processing phenomena under the common coding umbrella. In their view, the Theory of Event Coding (TEC) replaces "classical" approaches such as stage theory. Against this, I will propose that TEC may be viewed as complementary rather than contrasting in this regard. Second, some of the key concepts in TEC require clarification so as to become verifiable. The authors mention that TEC is only a loose conceptual framework – more a source of inspiration than a formal and verifiable theory. Stage and capacity notions about human information processing were not that elaborate either, but they at least contained a few basic predictions which could be – and have been – tested experimentally. For example, limited capacity theory failed to cover opposite effects in performance decrement in dual tasks (e.g. Navon & Gopher 1979; Sanders 1997). Is there any way in which TEC's key concepts can be shown to be wrong? Third, there is the question of the limits of common coding . Are there limits to the statement that perception and action-planning are "intimately related and sometimes even indistinghuisable processes?" (sect. 1, Introduction).

***1. TEC and stage theory.*** The authors mention Donders' (1868) original notion of decomposing choice reaction time (CRT) into a time taken by perceptual discrimination and a time taken by response selection – to use present-day terms – and the rediscovery of the stage concept during the late sixties. However, they fail to discuss how decomposition may be achieved. Thus, it is not conveyed that Donders' subtraction method already failed during the 1880s, leading at that time to the demise of the processing stage concept derived from subtraction. In contrast, Sternberg's (1969) additive factor method (AFM) has been remarkably successful not only in regard to categorisation of relations among the common variables affecting CRT, but equally to effects of energetical factors, such as psychotropic drugs and sleep loss (Sanders 1983; 1998, Ch. 9), and even to response intervals produced by animals in operant conditioning (Roberts 1987). Roberts has correctly noted that there are only a few instances in which a simple relation like perfect additivity generalizes over such a broad sprectrum of situations. The point of interest is that a conceptual notion will not live long without a proper methodology. Thus, Sanders (1983) never "proposed" four stages in a CRT task out of the blue; instead, the stages derived from solid evidence about additive relations among experimental variables. On the other hand, the AFM has never claimed universal validity; much research has actually aimed at delineating its domain (Sanders 1998, p. 69). Thus, it has been found to fail in cases of separable multidimensional stimuli (Miller 1982), Eriksen flanker stimuli (Ridderinkhof et al. 1995), and arousing stimuli (Sanders 1977).

Yet, these three cases are all interesting in their own right, providing new insights about old issues such as what defines a stimulus and routes of specific and aspecific processing. Again, issues of spatial localisation and, more generally, of selective attention have always been beyond the scope of the AFM. From the initial days of the "early-late" selection debate to Van der Heijden's (1992) postcategorical filtering, a stagewise processing notion has been implicit there. Yet, stage theory has added little for the simple reason that its prevailing methodology did not apply. The same can be said about the analysis of dual task performance. (Gopher & Sanders 1984). Without a proper experimental methodology a theoretical umbrella does not convince and, therefore, will not be of great influence.

The remarks in the target paper about the supposedly serial nature of the stages in CRT tasks are somewhat caricatural. First, serial procesing *within* a stage has never been proposed; and, second, it has been recognised ever since McClelland's (1979) simulations that continuous flow among stages is not necessarily at odds with the AFM (e.g., Miller 1988; Miller et al. 1995). Therefore, the issue of discrete versus continuous processing among stages is a separate research issue. In addition, the ineradicable bias persists that stage theory is merely data-driven; consequently, the widely discussed evidence for top-down control by way of, for example, motor preparation and expectancy is ignored, as are the ef-



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

fort and evaluation concepts in Sanders' (1983) stage model despite their clear reference to executive control and intentionality. Finally, the suggestion that stage theories merely distinguish between perceptual- and response-related processes is without foundation. On the contrary, the response selection stage has always been considered as a typical perceptual-motor interaction site. Stimulus-response compatibility plays a prominent role in the reviewed research on TEC and has also been traditionally a main variable in regard to response selection.

It is also interesting to note that, despite the rejection in the earlier parts, the stage concept seems not really at stake in the target article. For instance, two phases of stimulus encoding – that is, activation of feature bundles and their subsequent integration – are suggested which strongly reminds of the feature analysis and identification stages as derived from the AFM (Sanders 1998, p. 71) Again, TEC is said to apply to "'late' cognitive products of perceptual processing" and to "'early' cognitive antecedents of action" (sect. 1). According to stage theory, "late perceptual and early response products" typically reside in the response selection stage. The authors express doubts about the translation metaphor for response selection since it might imply an incommensurability between stimulus and response (sect. 1). However, in stage theory a translation metaphor has only been invoked in cases of *incompatible* stimulus-response relations, whereas joint perceptual-response codes have been commonly proposed for highly compatible reactions. Sanders (1967) suggested "elementary – that is, joint perceptual-motor – codes, evoked by any stimulus", and Greenwald (1970) emphasised ideomotor compatibility which seems fully consistent with a common coding notion. Hence, TEC and stage theory are complementary rather than competing. As the authors correctly note, though, stage theory has usually not had much to say about contents and structures involved in stages. A main merit of TEC may be in addressing contents and structures relating to response selection.

***2. TEC's conceptual elements.*** One of TEC's main concepts concerns intentionality, that is, goal-directedness of behavior. The target paper is fully correct that all performance studies presuppose instructions defining the intentions to be realised in an experiment. The paper is also correct that this is usually neglected in theorizing. According to TEC, intentionality renders perception and action-planning inherently similar and functionally equivalent. This is contrasted with the usual position of sensorimotor theories in which perception and action are usually separate realms connected by acquired arbitrary associations. The paper makes the useful distinction between distal versus proximal aspects of perception and action planning, the distal aspects relating to TEC's subject matter – that is, the late perceptual and early action processing. "Perception is actively acquiring information. Action would be blind without information of the sensory consequences."

Thus, it is not surprising that TEC also leans on direct perception views, such as affordances (Gibson 1977), to further underpin the essential role of action in perception, and vice versa. The authors mention Lotze's (1852) local signs as an example of an early classical predecessor. Perhaps the similarity is even more striking with von Weizäcker's (1940) *Gestaltkreis* – proclaiming the unity of perception and action – and with sensori-tonic field theory as proposed during the fifties by Wapner and Werner on the basis of results on the perceived vertical (Werner & Wapner 1952) – although the latter theory might lack the strong intentional overtone of the von Weizäcker formulation. For example, Derwort (1938) was among the first to note the relative constancy of the time needed to draw a specific figure, such as a circle, irrespective of its size. This result led to the far-reaching conclusion that specific forms are encoded in terms of the time needed for their production, which was taken as an expression of the unity of perception and action. The target article would probably not endorse this view which may be considered as an extreme example of an affordance, but it remains unclear what the authors actually mean by functional equivalence of codes for perception and action-planning. Identity? Stimulus and response codes are sometimes distinguished although on other occasions they are almost treated as identical. Do they mean comparable representations of perception and action in that both are based on features and integrated structures? Interaction? The authors refer to Adams' (1971) perceptual and motor trace in motor learning as an early example of functionally equivalent perception and action codes. Yet, Adams' traces cannot be said to be identical, the motor trace referring to a movement skill and the perceptual one serving as a reference for deciding whether motor movement had been successful. The two traces are interactive, though, in that both are a prerequisite for completing an action. Yet, in Adams' view, the perceptual trace was not "perceptual" in the usual sense, rather, it was a motor representation being reinforced each time a specific motor trace had been succesfully executed. In this way, Adams basically aimed at a neobehaviorist description of motor learning. In TEC the issue of learning remains somewhat obscure. Learning by reinforcement is avoided and TEC is more akin with perceptual learning as a process of differentiation and refining (Gibson 1969; Gibson & Levin 1975). This may be viable as long as dealing with, say, developing perceptual invariants, but has been much less succesful in more indirect skills such as learning how to read (Marr 1982).

The definition of an event is left open deliberately in TEC but, whatever the definition, a basic element of an event seems to be a bundle of feature codes, commonly shared by perception and action. Feature bundles need integration through feature binding. One way in which perceptual and response codes are thought to be controlled in a different way is that stimulus codes are manipulated by attention – through different weights of the features dependent on their relevance to the task – and response codes by intention – that is, by unfolding the planned action. The ample evidence for bottom-up parallel processing of elementary stimulus features (e.g., Sanders 1995; Theeuwes 1993) casts considerable doubt on the notion of attentional control of perceptual parameters – presumably with the exception of spatial localisation (e.g., van der Heijden 1992). However, the target article might consider features as dimensions in the sense of Kornblum et al. (1990), the degree of dimensional overlap between stimulus and response codes determining the degree of S-R compatibility. Without any dimensional overlap it is only possible to carry out a CRT task through laborious practice of associations between specific stimuli and specific responses.

***3. The limits of TEC.*** Some dimensions do coincide with "real" perceptual and response features, such as spatial localisation, movement direction, and optical flow. However, as argued by Kornblum et al. (1990), other dimensions might be quite more complex, which bears immediately on the direct-indirect issue. Imagine, for instance, a CRT task in which a stimulus consists of one of the visually presented digits 1–9 and the response of a pointing response to one out of 9 spatial positions on a horizontal row. This situation is not highly compatible but optimal compatibility is reached when the digits 1–9 correspond to spatial response positions in a regular order from left to right. Formally, digits and spatial positions have little dimensional overlap – the stimuli consisting of visual patterns and the responses referring to spatial locations. Yet, the common dimension is something like "quantity," which has been learned to regularly increase from 1 to 9 and spatially from left to right. Does "quantity" refer in TEC to a shared stimulus and response feature? This might well be argued but in that case the theory should incorporate laborious associative stimulus-response learning in order to assign the meaning of a certain quantity to a visual pattern, such as a "2" or a "6," and to assign quantities to spatial locations in a horizontal row from left to right. If quantity is considered to be among the common codes, TEC should place much more emphasis on cognition and learning. In addition, pattern recognition of the digit might not be included, despite the fact that identification of the pattern may be considered as the primary perceptual event. Actually, Kornblum et al. discuss dimensions in terms of the sensitivity of cross-dimensional scaling and not as primary features of stimuli and responses.



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

This leads naturally to the issue of the limits of TEC. As mentioned "quantity" may still be among the shared features of perception and action, and be described in terms of affordances. This is harder for some other CRT tasks. The target article deplores the fact that much research on CRT has been little related to action; at the best, the choice reaction was an unavoidable byproduct, the research on target classification (e.g., Sternberg 1975) serving as a prime example. It is undeniable that the human information processing tradition has been foremost interested in issues of perception, memory, and reasoning with, at least originally, little eye on action. Perceptual-motor interaction became a main domain for proponents of direct perception. The question, of course, is whether all information processing, direct as well as indirect, is relevant to TEC. Does TEC only refer to the domains of the ecologists or do perception and action planning also share a common representation in a Sternberg classification task? The issue, then, concerns the extent to and the conditions under which perceptual and motor features are equivalent, in the sense of strongly interacting. The authors review, among others, a wide variety of convincing studies on S-R compatibility in the spatial and in the movement domain, on imitation, on response-stimulus synchronisation, and on effects of an action on subsequent perceptual encoding. Most of the evidence, though, relates to spatio-temporal conditions in which functional equivalence is most likely to play a predominant role. Still, even in the above domain there are instances in which equivalence requires further clarification. Thus, the target paper discusses an example of equivalence in which perceiving a moving spot to the right corresponds to a movement to the right. So far so good. But there are also harder cases of movement compatibility for TEC, as exemplified by Warrick's (1947) rule. Normally a clockwise rotation of a dial is connected with an upward shift of a horizontal line indicator on a vertical display. However, this is the case only when the dial is positioned below or to the right of the vertical display. If the dial is positioned to the left of the vertical display, there is no clear association between clockwise rotation and an upward or downward movement of the display. What are the common codes relating a clockwise or counterclockwise rotation of a dial and the position of a dial to an expected upward and downward movement? Perhaps TEC could come up with a solution but, without further premises it might not be compelling.

To finish positively: TEC has been a source of inspiration to the Munich group, resulting in a series of beautiful studies. This may seem a sufficient rationale for TEC as umbrella concept. Yet, more needs to be done concerning delineation, specification, and limits. In addition attempts toward an all-compassing view should be avoided.

## Explanatory burdens and natural law: Invoking a field description of perception-action

Robert E. Shaw and Jeffrey B. Wagman

*Center for the Ecological Study of Perception and Action, University of Connecticut, Storrs, CT 06269.* **Robert.E.Shaw@uconn.edu**
**Jeffrey.B.Wagman@uconn.edu**
**http://ione.psy.uconn.edu/~cespaweb/people/wagman.html**

**Abstract:** Although we agree with **Hommel et al.** that perception and action refer to one another, we disagree that they do so via a code. Gibson (1966; 1979) attempted to frame perception-action as a field phenomenon rather than as a particle phenomenon. From such a perspective, perception and action are adjoint, mutually interacting through an information field, and codes are unnecessary.

We agree with **Hommel et al.** that any adequate theory of perception and action linkage should satisfy an intentionality condition – that perceiving refers to acting, and acting refers back to perceiving. Similarly, ecological psychologists generally agree that a *circular causality* holds between perceiving and acting, where agents perceive how to act to reach a goal and, then, the acting updates the next goal-specific perceiving, which then updates the next goal-relevant acting, and so on, until the goal is reached or the effort aborted. Goal-directed activities conform to a perceiving-acting "cycle" wherein information and control reciprocate under mutually shared intentions.

Different theorists may agree that perceiving-acting cycles exist but disagree on how they work. The authors suggest that the intentional link between perception and action can be adequately explained as a shared event "code." Ecological psychologists suggest, instead, that to use a nondynamical entity (e.g., a code) to explain a dynamical phenomenon is to improperly describe the phenomenon in question. A successful code, or program, is merely a time-free surrogate that must borrow its rate of unfolding from a yet unspecified dynamical process whose intrinsic timing actually runs the show (Shaw & Todd 1980).

Instead of a symbolic function (code), we propose a perception-action linkage via a tunable coupling between conjugate information and control fields – an expression of natural law (Kadar & Shaw 2000). Indeed the code only works when underwritten by such lawful dynamics. Here, what an agent sees is related to what it does by conjugate fields: Control paths (streamlines) in one field influence and are reciprocally influenced by goal-specific information (equipotential) curves in a conjugate field. This hypothesis is made plausible by the generalizability that field theory has shown over the past century.

The explanatory power of Newtonian mechanics was based on reductionism and mechanism. Ironically, this primary strength was also its primary weakness because reductionism renders some fundamental phenomena inexplicable. For example, from a Newtonian perspective, gravity seemed to require a mysterious "action-at-a distance" – a clear violation of proximity required by causal principle. By placing objects in a geometrodynamical field, Einstein provided a principled means for resolving this mystery. Fields, unlike particles, are continuously extended in space-time. By providing a mediating causal substrate, fields allow momentum transfers between bodies over arbitrary distances (Kadar & Shaw 2000). As a result of these properties, fields have surpassed particles as the primary objects in physics (Davies 1988).

This paradigm shift allowed great progress in the physical sciences, such as the successful field description of gravity and electromagnetism. This shift also promises to promote great progress in the life sciences. In traditional accounts, distal correlation ("action-at-a-distance") in biological processes (e.g., morphogenesis, embryogenesis, and neural growth and development) was attributed to a genetic code or developmental program. A field description of these processes, by contrast, suggests no code may be necessary. All such processes share the same basic sequence of constraints; as cells move, they exert forces that arise because of adhesive, chemical, or electrical gradients. These forces deform the surroundings constraining future cell migration patterns (Murray 1990; Turvey & Fitzpatrick 1993; see also Edelman 1996; Goodwin 1994; Purves 1988; Thompson 1942). This continual exchange of flows and forces promotes a circular causality in which the level of the collective both influences and is influenced by the level of the particle.

Artifactual explanations arise only if problems are framed improperly, such as field phenomena being mistaken for particle phenomena. Parsimony dictates that non-holonomic constraints (programs or symbols), should be used sparingly, and then only to set subtle boundaries on dynamical processes (Gould & Lewontin 1979; Kugler & Turvey 1987). Such constraints are expensive and inappropriate as organizational principles since Nature is extremely frugal.

Traditional perceptual theory follows a Newtonian analog. The fundamental separation between perceiver and perceived it assumes, creates an epistemological version of the "action-at-a-distance" problem (Kadar & Shaw 2000). As a consequence, an *a*



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

*posteriori* link must be forged between animal and environment and also between perception and action. By contrast, Gibson (1966; 1979) removed the need for such *a posteriori* links by making perception and action *adjoint,* capable of mutual and reciprocal interaction through an information field (Kadar & Shaw 2000; Shaw et al. 1992; 1995). For Gibson, perception and action are functions of organism-environment ecosystems where the fundamental problem is the perceptual control of action. Consequently, his paradigmatic case was a moving observer guided by a lawful dynamical relationship between perception and action. Just as moving particles create forces and gradients, so do moving observers. While electric currents generate magnetic fields, moving observers create velocity fields, that is, *optic flow fields.* Furthermore, this exchange of flows and forces promotes a circular causality in which egomotion generates a global flowfield and a flowfield creates the impression of egomotion (Kadar & Shaw 2000). In general, the perspective reflects several key trends in scientific explanation: (1) a shift from anatomical specificity to functional specificity; (2) distribution of representations; (3) the recognition of the importance of local exigencies and context dependency; and (4) decentralization of control.

Although codes are not equipped to explain the emergence of order in non-biological systems, they are frequently invoked to describe the same in biological systems (Goodwin 1994; Kelso 1995; Turvey & Fitzpatrick 1993). In standard accounts, perception and action relate indirectly via mechanisms (e.g., programs or codes) that seem to sit outside of natural law. Conversely, in field accounts, perception and action relate directly and lawfully. Furthermore, the burden of explaining the emergence of order by field-dynamics seems far less taxing than the burden of appealing to disjointed, discrete units such as codes (Weiss 1967).

Hence, parsimony dictates that the explanatory burden should be removed from the code construct and placed squarely on the shoulders of the field construct. Nothing of value would be lost from doing so, while the gain would be to move psychology, biology, and physics closer as natural sciences.

## Multi-level sensorimotor interactions

Stefan Vogt[a] and Heiko Hecht[b]

[a]*Department of Psychology, Lancaster University, Lancaster LA1 4YF, United Kingdom;* [b]*Man Vehicle Lab, Massachusetts Institute of Technology, Cambridge, MA 02139.* **S.Vogt@lancaster.ac.uk** **hecht@mit.edu** **http://mvl.mit.edu/AG/Heiko/**

**Abstract:** We share the authors' general approach to the study of perception and action, but rather than singling out a particular level of "late perceptual" and "early motor" processing for sensorimotor interactions, we argue that these can arise at multiple levels during action preparation and execution. Recent data on action-perception transfer are used to illustrate this perspective.

Our commentary takes an unusual perspective. We are fully sympathetic to the authors' plea for an integrated study of perception and action. We collaborated with them in their Munich laboratories, and the notion of common coding (e.g., Prinz 1992) has inspired our experimentation. Perceptual and motor processes typically interact with each other, and this interaction serves many disambiguating functions. We are not fully convinced, however, of the merits of singling out a particular level of "late perceptual" and "early motor" processing that should be privileged for code commonality and that in turn would leave "earlier" sensory and "later" motor processing behind as incommensurate. Instead, we argue for multi-level sensorimotor interactions that can arise at all perceptual and motor stages.

Take the example of planning to squeeze an orange for breakfast: we are happy to follow the Theory of Event Coding (TEC) in characterizing this event as being represented in a cognitive code that gets activated both during action planning and whilst perceiving a friend performing this action, and we are reassured that a common cortical area is involved in both situations ("mirror neurons" in lateral premotor cortex; Rizzolatti et al. 1996). However, it is equally clear that visuomotor processing does not stop at this point. Once the action of squeezing an orange is selected amongst a number of actions that were competing for execution, a new plethora of interacting visuo-motor processes becomes involved that guide the hand to the fruit, preshape its grip, place the fruit on the work surface, and so on. A network of parieto-frontal circuits has been shown to be involved in integrating the sensory and motor processes implicated in these activities (e.g., Burnod et al. 1999; Caminiti et al. 1998; Rizzolatti et al. 1998). Are we supposed to describe action selection in TEC terms and to adopt an entirely different perspective when studying visuo-motor processing of the latter kind, that is, during action execution? The cortical networks for action selection and execution are at least partially overlapping, and the principle that sensory and motor signals relevant for an action are processed simultaneously and interact at multiple levels certainly holds for both. Thus, whilst we acknowledge **Hommel et al.**'s battle against the dark ages of an isolated study of perception and action, we are concerned about their proposal of singling out a specific level of coding for sensorimotor interaction to arise. Instead, we take the latter to be a universal phenomenon.

For instance, whilst distal coding is clearly an important design principle, it is not the only medium in which vision and action can meet. Rather, egocentric frames of reference (or proximal coding) are typically used in visually guided actions. More generally, visuomotor interactions in the moment-to-moment guidance of action are, at present, documented in much greater detail (e.g., Jeannerod et al. 1995) than for the more complex processes of action selection. We do not see the former, rich body of evidence adequately reflected in TEC. Certainly, **Hommel et al.**'s proposal of an ensemble of abstract features (such as object size or location), allegedly used in perceptual identification and action planning, could not easily be stretched towards execution, which requires that objects be represented in partially different ways (semantic vs. pragmatic representations; Goodale 1997; Jeannerod 1997). Nevertheless, a theory of action planning would ideally address the well-documented behavioral dissociations between vision-for-perception and vision-for-action, in Milner and Goodale's (1995) terms, rather than leave us with a conceptual gap between planning and execution as TEC does.

We now review a recent study on action-perception transfer (Hecht et al. 2001) that was motivated by the common coding approach. It clearly demonstrates TEC's heuristic value, however, it also suggests that cross-modal interactions are just as tenable for explaining our findings as is a common coding account. Our example will also demonstrate some of the difficulties in determining whether two separate but interacting representations are formed, or whether one abstract representation is distilled from sensory and motor inputs.

The starting point for our studies was the recognition that effects of action on perceptual processing had so far only been described on the short time scale of concurrent interactions (e.g., Craighero et al. 1999) but not in the sense of lasting effects of motor learning on visual discrimination, which we coined "action-perception transfer" (APT). The two basic tasks used were a motor production task and a visual judgment task, and we predicted that motor practice without visual involvement should enhance visual discrimination due to an underlying common representational format of the two tasks. Experiment 1 in Hecht et al. (2001) employed a rhythmic timing task where participants either produced a two-cycle arm movement with a certain time ratio (e.g., 800 and 400 msec or 50%) that was verbally instructed in the beginning of each trial (motor task), or they had to name the time ratio of a two-cycle motion presented on a monitor (visual task). Twenty-one different ratios between 50% and 200% were used. Separate groups of subjects underwent pure motor practice or pure visual practice and received verbal feedback about their



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

movement or judgment. Subsequently, all groups were tested in motor and visual performance without feedback. This design allowed us to assess both transfer from perception to action (PAT) and APT. Indeed, we found transfer of practice in both directions, and APT in particular. That is, judgments in the visual test were more accurate in the motor practice group compared to a control group that received unrelated motor practice. Experiment 2 showed that APT did not depend on active movement but was just as pronounced for participants undergoing passive practice. Passive participants merely received kinesthetic feedback about the movement (their arms were fixed to a lever yoked to and driven by the active participant).

These results can certainly be explained in terms of common coding (or TEC), namely, in terms of an abstract code that represents features of the visual stimuli and arm movements, presumably their relative timing. However, it also fits the bill of interacting sensory and motor processes. Two reasons motivate us to entertain the latter alternative. Firstly, the former explanation is limited to timing tasks and would thus not cover the wider range of tasks for which APT may be expected. Secondly, during the visual test, TEC would leave no role for kinesthetic or motor coding. In contrast, a multi-level interaction account would allow matching of the novel visual stimuli to kinesthetic signals, which our participants had learned to interpret during motor or kinesthetic ("passive") practice. Furthermore, sensorimotor interaction allows motor encoding of the visual stimuli as mediator of visual-kinesthetic matching (visuomotor-kinesthetic matching). After evaluating the available evidence for each account, we tended to favor the latter, which would predict a network of visual, motor, and somatosensory areas simultaneously involved in coding the visual motion, rather than a single, amodal coding domain.

To conclude, the heuristic value of the common coding approach is undisputed, and our study exemplifies this. Rather than constructing common versus interacting representations as theoretical alternatives, one might envisage event coding to rely on an integrated network of sensory-dominant and motor-dominant brain areas, with event codes as emergent properties of this network. TEC's gap between action planning and execution, however, is a considerable price to pay for perception and action to meet, and it leaves quite a bit to be desired towards a fully integrated account of cognition and action.

## Perception and action planning: Getting it together

David A. Westwood[a] and Melvyn A. Goodale[b]

[a]*Department of Psychology, The University of Western Ontario, London, Ontario N6A 5C2 Canada;* [b]*Departments of Psychology and Physiology, The University of Western Ontario, London, Ontario N6A 5C2 Canada.*
**dwestwood@uwo.ca** **mgoodale@uwo.ca**
**http://www.sscl.uwo.ca/psychology/faculty/goodale/**

**Abstract: Hommel et al.** propose that high-level perception and action planning share a common representational domain, which facilitates the control of intentional actions. On the surface, this point of view appears quite different from an alternative account that suggests that "action" and "perception" are functionally and neurologically dissociable processes. But it is difficult to reconcile these apparently different perspectives, because Hommel et al. do not clearly specify what they mean by "perception" and "action planning." With respect to the visual control of action, a distinction must be made between conscious visual perception and unconscious visuomotor processing. Hommel et al. must also distinguish between the *what* and *how* aspects of action planning, that is, planning *what* to do versus planning *how* to do it.

**Hommel et al.** present a framework for characterizing the interface between "perception" on the one hand and "action planning" on the other. In their "Theory of Event Coding" (TEC), Hommel et al. suggest that a common medium (the "event" code) is used to represent both perceived features of the external environment (i.e., stimuli) and features of possible actions (i.e., responses). According to Hommel et al., action planning is accomplished at the same time that the target or goal stimulus is perceived, because the event code does not discriminate between stimulus features and action features. The TEC framework is both elegant and attractive because it suggests that action planning does not involve a discrete set of computational processes that must be described separately from perceptual processes. But this elegance and attractiveness is clearly superficial. The computational complexity involved in getting from sensory input to useful motor output is "hidden" in what the authors refer to as "early" perceptual processes, and "late" motor processes – processing stages that lie beyond the scope of the theory.

We have no argument with the notion that sensory and motor representations must be integrated for the planning and control of action. Indeed, how could it be otherwise? It remains an open question, however, as to whether or not the "event" coding scheme proposed by the authors adequately captures the nature of the interface between sensory processing and motor planning. In fact, it is very difficult to judge the merits of TEC because the theory as presented suffers from a problem that plagues "perception and action" research: a failure to provide an adequate definition of the critical terms "perception" and "action planning." This is not a trivial matter, because it has profound implications for the testability of the theory.

Nowhere in the paper do **Hommel et al.** state whether "perception" refers to conscious or unconscious processes. Indeed, at one point they declare that TEC is not intended to address the relationship between conscious perception and externally-directed movements such as pointing and grasping. One gets the sense that Hommel et al. do not consider this relationship to be of particular importance. This is rather puzzling in light of the substantial literature (for review, see Milner & Goodale 1995) suggesting that an important distinction can be drawn between the visual processes that lead to conscious visual perception and those that support the visual control of action. Surely any theory that purports to describe the interface between "perception" and "action planning" must address this issue, so that the theory can be considered in the light of existing neurological and behavioural evidence.

**Hommel et al.** are equally opaque about the notion of "action planning." At times, they seem to be talking about action planning as an issue of intention, as in deciding *what* to do; for example, deciding whether or not to pick up a cup of coffee. In this sense, the task is to select an appropriate course of action from a number of possible alternatives, given a particular visual environment (e.g., a table with cups of coffee and perhaps utensils) and a specific behavioural goal (e.g., to have a drink of coffee). At other times (e.g., in their discussion of attentional and intentional modulation of event coding), they seem to use the term "action planning" to refer to the problem of specifying *how* to accomplish an intended action, for example, planning where to place the fingers when grasping the cup of coffee. Indeed, their invocation of the visuomotor neurons in parietal cortex as examples of how "perception and action planning may be interfaced in the brain" suggests that they are focusing on the rather specific way in which visual inputs are transformed into the co-ordinate frames of particular effectors. But if TEC is to provide an account of how action planning unfolds, Hommel et al. must clearly distinguish between the *what* and *how* of planning.

The Milner and Goodale (1995) account of the two cortical visual systems is quite clear about the meaning of "perception" and "action." Perception refers to the conscious experience of visual sensation, the phenomenal awareness of the form, colour, shape, and position of visual objects in the world beyond our bodies. In this context, perception does not refer to all sensory processing, but to a specific type of processing that engenders conscious awareness of the external world. In the Milner and Goodale model, action refers specifically to the transformation of visual information into spatially calibrated motor outputs, which is of



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

course necessary to implement a particular course of action. The visual processes necessary for this *type* of action control (i.e., the *how* of action planning) are thought not to depend on or lead to conscious visual awareness; for this reason, the term "perception" is not applied to the visual mechanisms of action control. Of course, this does not mean that sensory processing is irrelevant for action control. Indeed, nothing could be farther from the truth; the control of action depends heavily on afferent (and reafferent) visual information, but the processing of visual information for action control is *fundamentally* different from the sensory processing that supports conscious visual perception. The processing is different because the action and perception systems are doing very different jobs. Perception delivers a detailed representation of the world that serves as a foundation for cognitive operations, in which the metrics of the world (with respect to the observer) are not well-specified; the visuomotor systems mediating action deliver accurate metrical information in the required egocentric coordinates, but these computations are spare and evanescent.

The Milner and Goodale (1995) model is often misinterpreted as suggesting that the ventral stream of conscious visual perception does not have a role to play in the planning and control of action. This is certainly not the case. Although the patient D. F., a visual form agnostic demonstrates remarkably good visuomotor abilities in some tasks, such as grasping objects and posting cards into oriented slots, she has tremendous difficulty with other aspects of "action planning." For instance, how could she possibly decide which of two visual objects to pick up when she cannot identify the objects? Moreover, how could she estimate the mass of a visual object for planning grasping forces, when mass is not directly specified by available visual information? Mass can only be determined by accessing stored information about familiar objects and materials (Goodale 1998), and this requires conscious visual perception to identify the object or the material out of which it is constructed!

In short, there are different types of "perception" and different types of "action planning." The distinctions that can be drawn between these are not insignificant. Since it is not clear to us which types of "perception" and "action planning" are addressed by TEC, it is too early to judge the relative merits of **Hommel et al.**'s proposal.

## How are events represented?

Gezinus Wolters[a] and Antonino Raffone[b]

[a]*Department of Psychology, Leiden University, 2300 RB Leiden, The Netherlands;* [b]*Department of Psychology, University of Sunderland, St. Peter's Campus, SR6 0DD Sunderland, United Kingdom.*
**wolters@fsw.leidenuniv.nl** **antonino.raffone@sunderland.ac.uk**

**Abstract:** We note two important problems in the Theory of Event Coding. First, because the representational format of features and events is unspecified, the theory is difficult to test. Second, the theory lacks a mechanism for the integration of features into an event code when the features are shared by different events. Possible ways of solving these problems are suggested.

According to the Theory of Event Coding (TEC), planning of actions is an off-line operation based on anticipated or intended action effects. There is much to be said in favour of a theory that integrates perception and planning and selection of actions in terms of anticipated perception of the effects. Elsewhere (Phaf & Wolters 1997), we have argued for such a process being realized in working memory. In that paper, working memory is conceptualized as a recent evolutionary development that created the possibility of using internal loops for holding and manipulating selectively attended representations. Such internal loops free an organism from the constraint that it can only react to physically present stimuli and that it can only form direct associations between objects and responses. The internalization of the sensorimotor arc of "perception-action-perception of results" allows for predicting the outcomes of actions without actually performing them, assuming that the present contents of the loops have access to the memory representations of prior events, prior actions, and prior action effects.

Although we endorse the general idea put forward in the Theory of Event Coding, we envisage a problem that questions not the validity but the usefulness of the theory. According to TEC, perceptual events and anticipated events underlying action planning are coded in the same representational medium. It is argued that all representations in this medium consist of sets of elementary features that can be activated independently, or glued together by a binding process to represent a coherent event.

An important question thus is: what are the elements of the common representational medium? TEC assumes these to be "events," that is, temporary integrated sets of features based on a neurophysiological binding process. The theory is rather unspecific, however, about what features are. Features can be simple elementary stimulus characteristics, like position, form, or color, but they can also be complex objects consisting of strongly associated complexes of simple features like geometrical figures, words, faces, or even scripts and schemas. This question is similar to the unsolved question of what "chunks" are in short-term memory research. As long as it cannot be made explicit *a priori* what exactly constitutes an actual or anticipated event, and which features are activated or integrated, mutual interactions between perception and action planning can only be determined post-hoc. One possibility to overcome this problem is that features may be defined in terms of their relevance for perceptual and motor categorizations. Thus, event properties which can be used for categorization are represented as features (Schyns et al. 1998; Ziessler 1990).

Another problem concerns the integration of events. According to TEC, features are coded by separate nodes, which are interacting through excitatory connections (see Figs. 1 and 2 of the target article). The active representation of events is generated by a two-stage process: activation of elementary features (individual nodes), and subsequent integration of these features into mutually segregated event representations by selective synchronization. However, the functional scheme described by **Hommel et al.** is not mechanistically plausible, since activation and integration cannot be easily separated in dynamical and neurophysiological terms. Specifically, the selective allocation of a shared feature (*f3*) to only one of two concurrently activated event representations, cannot be plausibly based on the oscillatory *synchronization within* and *desynchronization between* the two-event representations: since synchronization implies a transitive relationship, shared nodes may lead to interference-prone coding and readout of the events.

We suggest an alternative neural mechanism based on *activity competition,* which can coexist with synchronization coding. A simple network of units coding for five features was implemented, according to Figure 2 of the target article (*f1* and *f2* are allocated to represent Event 1, *f4* and *f5* to Event 2, and *f3* to both events, see our Fig. 1). A global inhibition was implemented, in terms of low-pass filtering of the average network activity. This global inhibition causes unit activity competition. Dynamic weighting of features was implemented in terms of a multiplicative input supplied to units coding for task-relevant features. Thus, at each iteration the net input from the external and from other connected units is multiplied by the weighting input. The product is then summed to the net input. This multiplicative computational scheme implements the attentional or preparational weights of event features, in terms of amplification (a weighting input equal to 0 means no amplification of the net input). Finally, an external input was supplied to one or more units.

As shown in our Figure 1A, an input-driven activation of the shared feature *f3,* in the presence of anticipatory weighting of feature *f2* allocated to Event 1, leads to a higher activation spreading



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

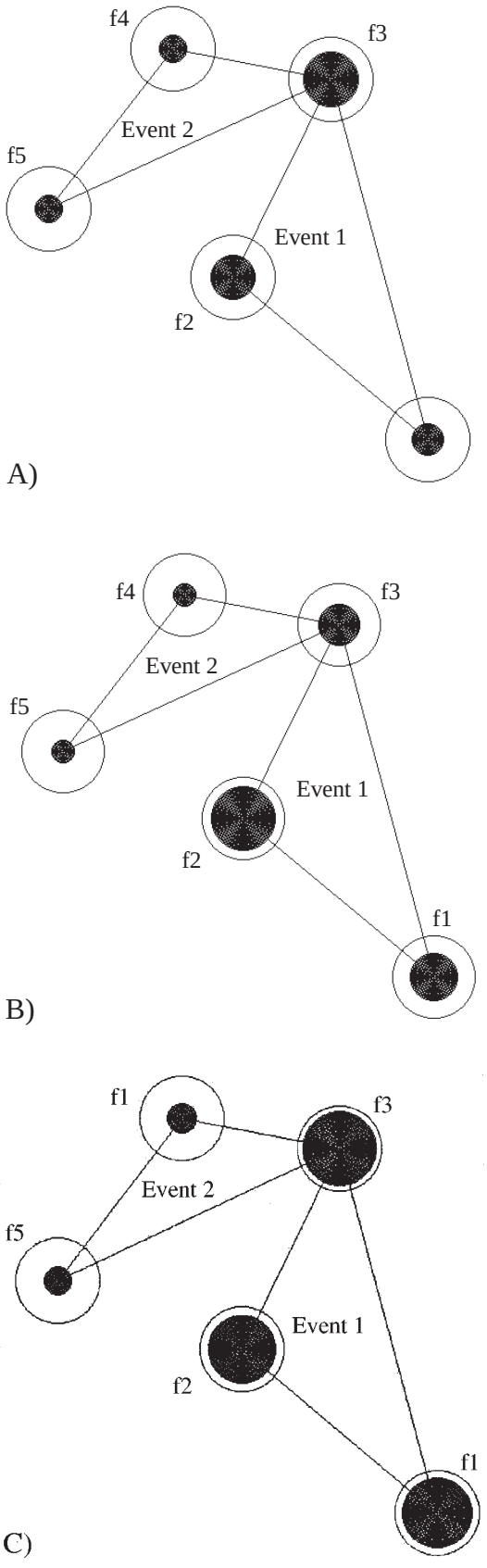



Figure 1 (Wolters & Raffone). Activation spreading and competition in a network of units coding for features of two events. The panels show the bi-directional connections and activation states after 500 iterations, in different conditions. The activation level is coded by the radius of the full circles. See text for more explanation.

to nodes coding for Event 1, without spreading to nodes coding for Event 2. Thus, activation spreading may be "directed" by feature weighting, even in the presence of shared features. In our Figure 1B, the input-driven activation of *f1* and *f2* (*f2* is also weighted), leads to a relatively unambiguous allocation of the shared feature *f3.* In our Figure 1C, *f1* is activated from external input and both features *f2* and *f3* are highly weighted: note how only nodes coding for the features of Event 1 are activated. Without global inhibition, all units are indiscriminately activated due to propagation through the shared unit.

As these examples show, activation participates in the integration process, which in turn biases the activation level of individual units. Therefore, activation and integration are not separate processes. Clearly, an additional synchronization mechanism is necessary for an unambiguous readout when multiple event representations are simultaneously active in the network. Such a synchronization can be accomplished in a parallel interplay with activation spreading and competition. Finally, in the case in which multiple event representations share features and get simultaneously activated due to their high weighting, a mechanism based on intermittent and graded (non-transitive) synchronization is necessary (Van Leeuwen et al. 1997).

## Scaling up from atomic to complex events

Jeffrey M. Zacks

*Washington University, Department of Psychology, Saint Louis, MO 63130.*
**jzacks@artsci.wustl.edu** **http://iac.wustl.edu/~jzacks**

**Abstract:** The Theory of Event Coding deals with brief events but has implications for longer, complex events, particularly goal-directed activities. Two of the theory's central claims are consistent with or assumed by theories of complex events. However, the claim that event codes arise from the rapid activation and integration of features presents challenges for scaling up to larger events.

**Hommel et al.** present a comprehensive framework for theorizing about interactions between perception and action (Theory of Event Coding, or TEC). As they note in section 3.2.5, this theory is focused on short timescales, up to a few seconds, but has intriguing parallels in research on larger-scale events. Given these parallels, it makes sense to examine closely the relationship between the current theory and extant theories of larger-scale event perception and action planning, and to ask how TEC might scale up to larger events. Of the universe of larger-scale events, goal-directed activities on the temporal scale of tens of seconds to a few hours appear to be particularly cognitively salient for cognition, much as medium-sized manipulable objects such as chairs and apples are salient (Barker & Wright 1954; Morris & Murphy 1990; Rifkin 1985). These are the mundane events people describe when asked what they did today: making coffee, fixing a flat tire, installing computer software.

TEC makes three basic claims (sect. 3.1), all of which are potentially applicable to everyday, goal-directed activities. First, perception and action share a common representational domain (*common coding*). Second, these codes refer to *distal features.* Third, an event code consists not just in the activation of a set of features, but also in their *integration* or binding. I will examine each of these in turn.

**Hommel et al.**'s exposition of the common coding claim closely follows the state of the literature on longer everyday events: In both cases, similar representations have been proposed for perception and action, but by different research communities. The clearest example concerns the part-subpart structure of activity (Zacks & Tversky 2001), which has been addressed independently by researchers studying event understanding and action planning. In narrative comprehension, event representations have been captured in event schemata (Rumelhart 1977), scripts (Schank &



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

Abelson 1977), and situation models (Zwaan & Radvansky 1998). All three postulate hierarchical part-subpart relationships governing the represented structure of activity. At the same time, researchers in planning have proposed hierarchical representations of action that reflect recursive goal-subgoal relationships (Newell & Simon 1972).

Recent perceptual studies have established that observers perceive event structure in terms of the same sort of hierarchical organization postulated for narrative understanding and memory (Zacks et al. 2001). It has been argued that this reflects the influence of shared representations guiding perception and action (Zacks & Tversky 2001). Such representations may be adaptive because information about actors' goals is correlated with distinctive physical features of activity, so a perceiver can better predict an actor's future activity by using a shared representation to mutually constrain perception of physical and intentional structure. In short, the common coding claim appears to have similar implications for larger scale events as for brief events, and those implications have begun to be explored.

Regarding the claim that perception/action representations refer to distal features of the environment, researchers studying larger-scale events appear to be ahead of the game. For speeded responses such as keypresses or simple reaching movements, and for simple stimulus patterns such as colored shapes or characters, the question looms large whether event representations are based on proximal features (muscle tensions, retinal stimulation) or distal features (movements of objects or changes of the state of a computer screen). However, I know of no theorists who have suggested that events on timescales longer than a few seconds are perceived or planned in terms of proximal features. Rather, the representations posited by researchers studying larger-scale events are uniformly described in terms of external objects, other people, and their relations – all distal properties. For example, in script theory (Schank & Abelson 1977) the primitive operators (Schank 1972) include physical transfers of objects, ingestion, and speaking, all of which are underdetermined relative to proximal features.

**Hommel et al.** present two arguments for the use of distal representations (sect. 5.2). First, action planning based on proximal features is inefficient. Second, prediction of future stimulus input is easier with distal features. They present a range of evidence in support of these arguments for small-scale events. This evidence is crucial because for these sorts of events the arguments for distal features, though reasonable, are not overpowering. However, for complex events the limits of proximal features become painfully clear: Although it may be plausible that the system plans a ballistic reaching motion in terms of the individual muscle contractions involved, planning a trip to the doctor's office is another story.

Thus, the implications of the TEC view for complex events have been more or less assumed in the literature. This is apparently because the arguments for distal features become overwhelming as they scale up.

The final claim of TEC is that the formation of an event code consists of two discrete stages: In the first stage features are activated and in the second stage they are integrated. The first stage facilitates processing of other events with overlapping features, whereas the second stage interferes with processing such events (sect. 3.2.2). Importantly, these stages have an intrinsic dynamic, which unfolds rapidly (sect. 4.1).

There is some evidence that representations of larger-scale events can prime each other. In one paradigm, participants studied a series of short stories that contained pairs of stories with overlapping features. They then answered questions about these stories. Under some study and test conditions, answering a question about a story was facilitated when it was preceded by a question about a different story that shared thematic (Seifert et al. 1986) or structural (McKoon et al. 1989) features. This demonstrates the possibility of feature activation facilitating processing of related events (the first stage in TEC) for complex events. Inhibition of related events (the second stage) has also been demonstrated for complex events. Radvansky (1999) had participants study a series of sentences about a number of events, which were distinguishable by their different locations. Participants then made speeded recognition judgments about the sentences. Responses to test sentences were slower when the previous sentence had referred to a different event that shared an object with the event of the test sentence.

These two sets of findings show that shared features can lead to both facilitation and interference for complex events as well as for simple ones. On the surface, this suggests that the activation/binding component of TEC may scale up. However, in these studies facilitation was observed more often, either with equivalent (McKoon et al. 1989; Seifert et al. 1986) or much longer (Seifert et al. 1986) delays, than inhibition was (Radvansky 1999). Moreover, perceiving and acting in larger-scale events unfolds over a much longer timescale than the one postulated for the two stages in TEC. This suggests that although the activation-plus-integration notion may apply to events on longer timescales, an important modification will be needed to scale up. On longer timescales, the intrinsic dynamics of automatic activation and integration are probably less important than constraints imposed by the task being performed and the semantics of the particular events.

In short, the first two of TEC's claims are consistent with theories of larger-scale events: Common coding is implicit in those theories and has begun to be explored directly, and distal features are assumed. However, the activation-plus-integration claim will likely need modification to scale up.

# Authors' Response

## Codes and their vicissitudes[1]

Bernhard Hommel,[a,b] Jochen Müsseler,[b]
Gisa Aschersleben,[b] and Wolfgang Prinz[b]

[a]*University of Leiden, Section of Experimental and Theoretical Psychology, 2300 RB Leiden, The Netherlands;* [b]*Max Planck Institute for Psychological Research, D-80799 Munich, Germany.*

**{muesseler;aschersleben;prinz}@mpipf-muenchen.mpg.de**
**www.mpipf-muenchen.mpg.de/~prinz hommel@fsw.leideniniv.nl**

**Abstract:** First, we discuss issues raised with respect to the Theory of Event Coding (TEC)'s scope, that is, its limitations and possible extensions. Then, we address the issue of specificity, that is, the widespread concern that TEC is too unspecified and, therefore, too vague in a number of important respects. Finally, we elaborate on our views about TEC's relations to other important frameworks and approaches in the field like stages models, ecological approaches, and the two-visual-pathways model.

### R0. Introduction

As we stress in the target article, the Theory of Event Coding (TEC) is meant to be a broad framework for understanding relationships between perception, cognition, and action planning, not a specific model or theory. Accordingly, and not surprisingly, a number of commentaries address basic theoretical and methodological issues regarding the nature and appropriateness of that framework *in toto,* as well as its relation to other frameworks and approaches. Others raise more specific issues and offer detailed suggestions for extensions, modifications, and so on. We have found most of the commentaries helpful for shaping what TEC is meant



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

to be and what it's not. Our reply is organized into three main sections, the first dealing with TEC's scope (R1), the second with TEC's degree of specificity (R2), and the third with TEC's relation to other frameworks and approaches to perception and action (R3).

## R1. The scope of TEC

### *R1.1. Limitations*

There is widespread concern among a number of commentators that the approach we take is too narrow (if not narrow-minded), given the enormous complexity and richness of the phenomena to be explained, – that is, the ubiquity of tight couplings between perception and action that can be seen in all animals in their natural settings, – and given the intricacies of the dynamical changes that these couplings undergo in the stream of ongoing behavior. This concern comes in two flavors: Some claim that the *approach* we take is flawed in principle. Others claim that TEC's *coverage* of pertinent phenomena is arbitrary and far from complete.

**R1.1.1. Approach.** According to some commentaries, TEC is flawed in principle because, as far as theory goes, it still maintains a dualistic view of perception and action and, as far as methodology goes, it strongly relies on experimental paradigms with arbitrary assignments of isolated responses to isolated stimuli. With these characteristics, the argument goes, there is no way to adequately capture the various natural forms of mutual interaction between perception and action, or to do justice to the circular causality inherent in animals for acting and perceiving.

**Shaw & Wagman** emphasize that perceiving and acting unfold in circular causality over time. From this perspective, they criticize a concept of code which is central to TEC as "a time-free surrogate that must borrow its rate of unfolding from an yet unspecified dynamical process." Instead, with a side-view on physics, they argue for a field perspective for research on perception-action cycles, as proposed by Gibson (1966; 1979). Like physics, which has long been moving from (Newtonian) particles to (Einsteinian) fields as its primary objects of inquiry, research on perception and action should move ahead from shared codes to shared information fields in organism-environment ecosystems. A move along this line, they argue, would relieve us from both dualisms: perception/action and animal/environment.

A similar point is raised by **Richardson & Michaels** and by **Kim & Effken**. They both argue that the issue of linking perception to action may perhaps be a problem for the minds of certain (conceptually misguided) cognitive scientists, but definitely not for the minds of the animals under study. These animals and their ancestors would not have survived if they had not been furnished with tight and efficient couplings between perception and action from the outset. Hence, dualism does not exist in nature but only in our theories (which speak of codes for perception and action), and in our experiments (which require arbitrary mappings between isolated, particle-like stimuli and responses). The common call of these commentaries is to abandon dualisms and codes altogether and to adopt an ecological view that treats the animal and its environment as a single system, whose performance is analyzed in terms of notions like information, information fields, and affordances.

In the same vein, TEC is criticized for not being explicit on how the alleged representations (i.e., event codes) are individuated and grounded: in perception or action (**Galantucci et al.**), in information or reality (**Richardson & Michaels**). Instead, since "there is no event ontology and no information specifying these events," TEC must put all of the explanatory burden on the perceiver/actor's mental operations and their "old-fashioned associationistic" processes (Richardson & Michaels).

Vis-à-vis these critical objections, we try to delineate once more what TEC is meant to be and what it is not.

*First.* TEC focuses on perception and action planning in humans. TEC's functional architecture is meant to account for the operations involved in action selection and action preparation. In any case, we believe that this architecture is functional in humans. We have no speculations to offer at which point in evolution this architecture emerged.

*Second.* According to its thematic focus, TEC requires a methodological approach that allows one to study human perception and action planning under experimentally controlled conditions. In this context, we wish to defend the often criticized reaction-time tasks that are, in fact, paradigmatic for the evidence on which much of TEC is grounded. Depending on perspective, this task is criticized (or even exposed to ridicule) for being both too simple and too difficult. On the surface, pressing a key is, of course, an extremely simple action in terms of the spatiotemporal coordinations required for its execution. Hence, when compared to more natural and more complex interactions with the environment, key-presses appear to be extremely simple, if not simplified actions (cf., e.g., **Pisella et al.**). However, as other commentators point out, pressing a particular key in response to a certain stimulus on the basis of a previously agreed-upon, arbitrary mapping rule, is, at the same time, also a highly unnatural and artificial task to perform (cf., e.g., **Galantucci et al.**). Humans can do it on the spot, but monkeys have a hard time learning it. Still, despite these objections, we hold that the choice reaction-time task captures exactly the type of performance TEC is meant to account for, that is, the linking of perception with the planning and preparation of arbitrarily selected action. As we point out below (in sect. R2.4), such performance may appear to be unnatural from an evolutionary point of view, but it is certainly not unnatural to all of us every day.

*Third.* From these limitations in scope and method we can gather what TEC is *not* meant to be. TEC is not meant to account for the online interplay between perception and action. In other words, TEC does not speak to issues such as spatial and temporal coordination of actions and environmental events, or to the fine-grained time course of speech and language processing, as some suggest it should (**Pisella et al.; Galantucci et al**.). Of course, we agree that it is often not easy to delineate (offline) planning from (online) interaction, but we do believe that these two functions need to be distinguished (see sect. R2.2 and Fig. 1).

*Fourth.* According to its scope and mission, TEC does rely on dualism in method, but certainly not in theory. We cannot see what is wrong about methodological dualism, when the goal is to provide a functional analysis of how actions are selected under given circumstances. In order to achieve this goal, we need to individuate possible circumstances (in the environment) and possible actions (in the person), and we need to distinguish between the percep-



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

tion of these circumstances and the planning and execution of these actions. It may be true that these instances can never be so neatly individuated in natural perception-action cycles, but why should one worry about creating an artificial experimental situation for the sake of a functional analysis one wishes to perform? What we certainly need to do is resist the temptation of importing dualism from method into theory. This is exactly what TEC tries to achieve: avoid dualism in theory while recognizing the need for a dualistic approach in method.

*In summary,* TEC's scope is narrower than some of the commentators are implying. TEC is not a comprehensive framework for perception and action in general, but rather a specific framework for the cognitive basis of perception and action planning in humans. TEC's mission is to provide a functional analysis of the underlying processing architecture, not an ecological analysis of complex interactions between the perceiver/actor and his/her natural environment. Hence, if one envisages the broad scientific enterprise of developing a more comprehensive theory of relationships between perception and action in animals (e.g., **Hochberg; Cisek & Kalaska**), TEC may become part of this enterprise – co-existing with other frameworks that address other issues. We hold that it is both legitimate and productive to detach this part from the rest of the enterprise. Due to its enormous flexibility, human action is special, and it can be studied with a special methodology that cannot easily be applied to animals. It is, of course, an open question to what extent human action also calls for a special theory.

**R1.1.2. Coverage.** Some other commentators who are more sympathetic with our general approach still maintain that TEC's coverage of the perception-action domain is arbitrary and incomplete.

*Early perception/late action.* Some criticize TEC's deliberate silence on what we call early perception and late action, that is, on low-level operations on the input and output side. One critical argument is that important issues remain unaddressed when low-level operations are disregarded. For instance, **Vogt & Hecht** criticize that TEC leaves an explanatory gap between action planning and execution. A similar point is raised by **Westwood & Goodale** in a more general sense. They argue that "the computational complexity involved in getting from sensory input to useful motor output is "hidden" in what the authors refer to as "early" perceptual processes and "late" motor processes," and that, since TEC does not address these processes, it is rated as "clearly superficial."

Another criticism is that, by focusing on late perception and early action, TEC fails to recognize the fact that perception and action are no less tightly entwined at lower levels of processing and, therefore, misses the opportunity to extend its coverage to also include early perception and late action. Accordingly, a number of supplements to TEC are suggested. **Cisek & Kalaska** argue that TEC's emphasis on the perceptual consequences of action is certainly not unique to advance flexible action planning. Rather, the principle of predictive feedback control is already abundant in simple animals and their situated activity. Hence, the structures envisaged by TEC must, in evolutionary terms, be considered recent specializations of very old principles. Likewise, **Bryson**, after discussing evidence on neurons that reference common feature maps for action perception and action control, stresses "that perception and action are probably unified in a number of ways of which TEC is but one." A similar point is made by **Dinse** from a neurophysiological and neuroanatomical perspective. Based on work with modified action and on anatomical evidence, he emphasizes the importance of crosstalk, feedback connections, and strong interactions at many levels in sensory and motor streams, as well as the role of temporal overlap between low-level and high-level processing. On this evidence, he argues that common representations for perception and action are probably not limited to higher processing levels and brain areas. In a similar vein, **Vogt & Hecht** discuss a study whose results can be accounted for in terms of both high-level event codes and low-level sensorimotor interactions. In their view, TEC should be expanded to (or even be replaced by) a multi-level interaction framework that relies "on an integrated network of sensory-dominant and motor-dominant brain areas, with event codes as emergent properties of this network."

We have three comments to offer in reply to these challenges and suggestions. *First,* we would like to reiterate what was said above: TEC is deliberately specialized and selective in that it focuses on the cognitive underpinnings of human action planning. *Second,* we certainly admit that our target article is sometimes less clear about the notion of action planning than it should have been. As **Westwood & Goodale** point out, we were not always consistent in keeping what they call the "What" and the "How" of action planning and execution as separate as we surely should have done. *Third,* and most importantly, we are at this point not convinced that it would be wise to broaden TEC's framework as suggested. As said above, we do of course agree that there is much more to action representation and control than what TEC has to offer. We, too, take it for granted that tight couplings and strong interactions between perception and action are ubiquitous in the animal kingdom, and that they are, at least in higher animals, implemented at various levels of coding. We maintain, however, that strong interactions and tight couplings do not necessarily imply common codes for perception and action. Consider, for instance, a population of neurons in the motor cortex that code for both certain movement sequences and their triggering conditions (cf. **Bryson**'s discussion of Graziano's work). Such neurons may be said to embody tight couplings between input and output without, however, providing common codes for them. Common coding would only emerge in a system of such neurons if a similarity relation holds between the triggering conditions and the movement sequences they code for (as the case of mirror neurons seems to suggest). Hence, common coding is a special form of tight coupling – one that allows for mappings of input to output (and vice versa) *by virtue of similarity.* This is why we, for the time being, resist the temptation of expanding TEC as suggested. Since TEC relies on the notion of common coding, it applies to representational systems whose input and output share the same coding dimensions.

*Stating facts/directing actions.* A different, though related supplement is offered by **Millikan,** suggesting a more precise notion of (mental) representations. We admit that the target article is quite vague on this issue. Nevertheless, we are not sure what Millikan's suggestions lead to. If one considers, as she points out, mental representations as entities that have representing as their (proper) function, two questions emerge: One is, what the function of represent-


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

ing means for the representations. Millikan seems to suggest that there is some "system" that uses and interprets them. What system? What use? The second question is what is represented. Millikan offers what may be termed a dual-face view of representations: they state facts (about distal affairs) and they direct action (in accordance with these affairs). This dual-face nature of representations, she argues, must have emerged in the early evolution of animals – in any case much earlier than TEC posits. This view seems to imply that directing action cannot occur without stating facts about distal affairs. We hesitate to subscribe to this principle. Why should it not be possible that sensory input gets translated into motor output without any reference to distal affairs?

### *R1.2. Extensions*

So far we have emphasized TEC's self-imposed limitations and have argued against a number of suggested extensions which we feel are inappropriate or at least premature at this point. In this section, we discuss other suggested extensions that we feel can in fact help to broaden the range of domains to which TEC's reasoning can be applied. We welcome these suggestions and invite further elaboration. We will go through the domains for the proposed extensions in order, beginning with those in TEC's close vicinity, proceeding to more remote fields.

**R1.2.1. Neuroimaging of action.** TEC outlines a functional architecture for perception and action planning without offering speculations about its implementation in the brain. Important extensions in this direction are provided in the brief review of recent neuroimaging studies on action perception and production provided by **Chaminade & Decety**. It appears that at this point two major conclusions can be drawn from this work. One is that natural action is specialized – in the sense that different brain structures are involved in the representation of biological and non-biological motion. The other is that the perception of natural action is supported by the same brain structures that are also involved in the generation of those actions (the premotor and parietal cortical areas). As **Chaminade & Decety** point out, these findings give strong support and provide important extensions to some of TEC's central claims. As we briefly discuss in the target article, related evidence has, over the past decade, also been accumulated in electrophysiological and TMS studies (Gallese et al. 2002; Jellema & Perrett, 2002; Rizzolatti et al. 2001). This evidence has been taken to indicate the existence of a *mirror system* in the brain that may work in two directions. One, where the perception of action is constrained by the perceiver's own action competencies. The other, where the execution of action may likewise be constrained by the actor's own perceptual experiences. The first direction is in line with motor theories of perception, the second in line with perceptual theories of action (cf., **Vogt & Hecht**).

**R1.2.2. Anticipation and intention.** TEC believes in a crucial role for the representation of action effects in action planning, and this applies to the representation of both expected and desired outcomes of actions (anticipation and intention, respectively). As **Hochberg** points out, this view has many predecessors. Accordingly, it isn't surprising that it gets broad support, partly from a systems-control perspective (e.g., **Cisek & Kalaska; Olivetti Belardinelli & Basso**), partly from an ecological perspective (e.g., **Kim & Effken**), and partly from a metaperspective that believes (or, hopes) that TEC may help to combine the systems control and ecological approaches (**Jordan**). Beyond this general support, **Rosenbaum** offers a specific computational demonstration of anticipation in action planning. Discussing the relative merits of forward and inverse approaches in motor planning, he points out that for both approaches a strong reliance on anticipation of the consequences of prospective motor acts is indispensable. Further, he provides a computational demonstration of end-state anticipation as a means of dealing with the redundancy problem in motor control. A demonstration like this suggests that TEC can be extended, as claimed in other commentaries (e.g., **Pisella et al.**), to also include "later" motor stages of action planning and become computationally more specific.

**R1.2.3. Attention. Ivanoff & Klein** address TEC's possible contributions to help clarify, or perhaps solve, a long-standing issue in the domain of visual attention. This issue refers to the role of (oculo-)motor factors for the orienting of attention in the visual field (Klein's oculomotor readiness hypothesis and Rizzolatti's premotor theory of attention; see Klein 1980; Klein & Pontefract 1994; Rizzolatti et al. 1987). As they point out, the evidence is ambiguous so far. A number of studies do support the claim that attention is driven by motor intentions, whereas others do not. As Ivanoff & Klein point out, TEC could perhaps inspire a solution to this conflict, because, rather than identifying attention with intention (as premotor theory does), TEC considers them to be two closely related, but still distinct processes. The closeness of this relationship might well depend on task settings, that is, on how closely the motor and the attentional tasks are coupled. This is certainly an interesting suggestion which could stimulate further research in this field.

**R1.2.4. Language.** Our insistence that TEC is not conceived to cover skills involved in the perception and production of spoken and written language does not, of course, imply that we are unwilling to follow invitations into this domain. For example, **Hochberg** invites us to do so because he believes that language provides prime examples of closely entwined perception and production skills – not only at the level of phonemes and syllables (as the motor theory of speech perception suggests), but also at the level of text and discourse. **Galantucci et al.**'s discussion of the motor theory of speech perception is certainly not meant as an extension to TEC but, rather, as an alternative approach from which one can learn why TEC is flawed. Still, as said above, though TEC is so far not conceived to address language processing, we are certainly open to extensions in this domain, too. One thing that we find attractive about motor theories of perception in general (i.e., not only in the speech domain, but also in domains like music, action, and attention), is their potential for offering a solution to the grounding problem, that is, the grounding of perception in action. In this regard, we certainly agree with one of **Galantucci et al.**'s key arguments.

Further, a strong case is made by **Hartsuiker & Pickering**, who claim that natural language processing is governed by theoretical principles similar to the ones TEC of-



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

fers for perception/action in general. They argue that language should not be excluded from TEC's scope for two reasons. One is that natural communication proceeds in dialogues, where one's speaking and one's listening to others tends to be closely linked and tightly coupled (not to mention one's listening to one's own speaking!). The other is that there is ample evidence for shared representations for comprehension and production in the language domain, as well as a crucial role of these shared representations in language-based communication and understanding. In sum, our target article is perhaps somewhat too much on the defensive vis-à-vis language and language processing, for this field seems to allow attractive extensions of TEC.

**R1.2.5. Complex events.** So far, TEC has mainly considered short-lived, particle-like events such as arrows or circles that come and go on a screen, or hands and fingers that go up and down on a key pad. Obviously, this is a serious limitation which needs to be overcome. Some commentators suggest pertinent extensions.

One such extension is suggested by **Lane et al.**, who use their CHREST model of active perception to demonstrate that the logic inherent in TEC's assumptions can also be applied to the sequential organization of active perception – such that the input information available at a given time is used to compute an output which, when executed, alters the input, and so on. CHREST seems to be related to TEC in at least two aspects: (1) action outcomes play a crucial role in both; (2) the same format is used for input and output representation. But one of the major differences is that CHREST is much more explicit about memory structures than TEC is (not to mention that CHREST is computationally much more specific). The interesting point here is not to compare one with the other but, rather, to realize that basic assumptions of TEC prove to be useful in a computational approach dealing with the perceptual exploration of complex scenes and events.

A similar point is raised by **Chown et al.** They take a look at TEC from the broader perspective of cognitive maps for navigation (their PLAN model). From this perspective, they argue, TEC needs to be extended in two important ways. One is that, since TEC fails to capture sequences of perception-action cycles as cognitive maps like PLAN do, TEC needs to go beyond individual cycles. Second, a mechanism for perceptual learning needs to be incorporated in TEC, perhaps based on Hebbian learning. This mechanism should be capable of generating more flexible, prototypic event codes than TEC's present scheme concerning feature combinations and abstractions allows for. Again, the point is not to compare the two approaches but rather to demonstrate that they are compatible with each other.

**Zacks** examines TEC from a still more remote perspective, namely, the perspective of research on the representational underpinnings of (relatively) large-scale events in everyday life like, for instance, making coffee, fixing a flat tire, or installing computer software. He finds much commonality between TEC and this research, for example, regarding crucial principles like common representations for perception and action, and reference to distal affairs. Remarkably, there appears to be no serious alternative to these two theoretical principles in this domain of study. Accordingly, there is an interesting lesson to be learnt from this comparison: The seemingly simple key-pressing tasks that support TEC can be regarded as down-scaled versions of tasks involving more natural, large-scale events. We take this as support for our claim that the traditional response-to-stimulus-mapping view should be replaced by an event-representation view.

However, as **Zacks** points out, there are also limitations to the parallels. In the processing of large-scale events, there is no equivalent to the pattern of activation-plus-integration that TEC suggests. We don't find this too surprising. The scheme of activation/integration should, in our view, be considered a short-lived automatic consequence of the presentation of brief stimuli – a sequence of processes that operates on a small time scale. For large-scale events that are extended over minutes and hours, we do not see anything equivalent.

## R2. Specificity

A further recurrent theme in a number of commentaries is that TEC is underspecified and, hence – at least in its present form – not testable and falsifiable. **Hochberg** and **Sanders** raise this issue in a general sense. Others, like **Chaminade & Decety**, **Pisella et al.,** and **Westwood & Goodale**, criticize its underspecification with respect to presumably involved brain structures. More specific aspects of underspecification are addressed by **Oriet et al., Shaw & Wagman**, and **Wolters & Raffone**.

Yes, we agree that TEC is underspecified in many ways and, as we note in many places in the target article, it is deliberately so. Before going into details, we would like to say a few words about the virtue of vagueness in science – with special reference to TEC. We believe in this virtue for two reasons.

The first reason has to do with TEC's theoretical mission vis-à-vis the dominant traditional views in the field which treat perception and action as two more or less separate functions. One of TEC's central messages is that this view is mistaken and must be replaced by a new framework – as we outline it. Accordingly, TEC's main mission at this point is to stimulate deliberations and discussions about basic principles of perception/action architectures. We hold that global principles should be clarified before local theories are made. In fact, we see the world of cognitive science populated with too much precipitate overspecification in local models and theories, whose underlying global principles have not been discussed and clarified before.

The second reason has to do with TEC's strategic mission. As we state explicitly in the target article, we place it much more in the context of discovery and exploration than in the context of testing and falsification. Hence, its strategic goal is not only to stimulate discussion of theoretical principles, but also to act as a heuristic tool for stimulating new research and inviting new extensions and specifications. In the context of discovery and exploration, underspecification is a heuristic virtue, but overspecification is a deadly sin: Underspecified frameworks can act as sources of inspiration for new ideas and new research, whereas overspecified theories are bound to fall into oblivion.

However, we do not mean to imply that the goal of science is underspecification. In the following we shall go through some of TEC's central concepts and discuss a number of specifications suggested in the commentaries.


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

### *R2.1. Perception*

Some commentators focus on the notion of perception as it is used in the TEC framework and how it could, or should, be used in a broader sense.

**R2.1.1. Perception, action, and intention.** It has often been claimed that the proper function of perception is not only to state facts about distal affairs but also to direct forthcoming action (as **Millikan** puts it so elegantly). Traditionally, due to their roots in epistemology, theories have emphasized the representational function of perceptual systems, that is, their role in stating facts. However, from time to time, their action-direction potential has been emphasized, too. More than a century ago, motor theories of perception were the first to emphasize the role of motor representations and, accordingly, the action-directing power of perception (Scheerer 1984). More recently, motor theories of perception have gained support in domains like speech perception (Liberman & Mattingley 1985; **Galantucci et al.**), and movement perception (**Chaminade & Decety**; Prinz 2002). A similar perspective is entailed in Gibson's notion of affordances, that is, information specifying the action-directing potential inherent in a given stimulus (Gibson 1966; 1979). In a similar vein, the action-directing capabilities of perception have recently become emphasized from an evolutionary point of view (e.g., **Cisek & Kalaska; Galantucci et al.**): Obviously, selective pressure has formed perceptual systems to optimize their capacity for directing overt action – in any case, more than their covert capacity for stating facts. Hence, one may argue that their proper function is much more related to the directing of action than to the stating of facts.

TEC is certainly sympathetic with this general perspective, but at the same time it goes one step beyond. It believes that perception may lead to, or often imply, intention and action, but TEC also stresses the fact that perception is *preceded by* intention, that is, that perception is inherently intentional by itself. In everyday life perceptual activities are always embedded in the dynamics of the perceiver's intentional situation, and so it is in any experimental setting. In each and every experiment, instructions come first and only then comes a stimulus that leads to a particular response according to instructions. However, theories of task performance tend to commence with the stimulus and forget about instructions (Broadbent 1993; Prinz 1997b). These theories fail to acknowledge the fact that the stimulus is always perceived with reference to the pre-established intentional state. As **Jordan** points out, TEC does not only acknowledge this fact but also offers a mechanism to account for the impact of intentional states on attentional selection in perception.

**R2.1.2. Perception and awareness.** Another burden from the epistemological heritage is that the notion of perception tends to go along with the notion of awareness, as **Westwood & Goodale** suggest. This is, of course, a heavy issue with deep philosophical implications, and space does not permit us to go into a principled discussion. The only thing we can offer is a pragmatic remark. The way TEC speaks about perception does not entail the notion that perceptual processing goes along with awareness. We hold that awareness may, under certain functional conditions, *emerge* in the course of perceptual processing, or, perhaps, as a result of it. At this time, we do not understand what these functional conditions are, and future theories of perception will have to work on identifying them. We hold that the criterion of awareness can, at best, be indicative of those (yet unknown) functional conditions. In other words, awareness can sometimes emerge in perception (as it can in any other cognitive function) without, however, playing a functional role in itself.

In our view, theories that believe in such a role make the mistake of taking method for theory. Naturally, the criterion of awareness (i.e., availability for verbal report) plays an enormously important methodological role in research with human participants. However, the importance of that methodological criterion must not be confused with the importance of that factor in theory. In fact, we are not dualists enough to believe that awareness can, in itself, play a role in perceptual processing. For instance, there is ample evidence of perception without awareness in a number of tasks that require perceptual identification (which, according to **Westwood & Goodale**, must rely on processing in the ventral stream; see e.g., Klotz & Neumann 1999). Hence, it seems that perceptual identification can be efficient in the absence of awareness.

### *R2.2. Action planning*

Instead of addressing all the processes that bring about an action, TEC focuses on what we call "action planning." In our understanding, the term refers to processes that prepare the system to reach a particular goal, that is, to produce an intended effect. This preparatory function has three important implications. First, action planning needs to precede the action or action element that is planned. Under tight time constraints, as in typical reaction time experiments, planning, and execution may go hand in hand, that is, what is planned is carried out as soon as planning is completed or at least sufficiently progressed. Yet, in daily life many actions will be planned to some degree some time before the conditions for execution materialize. Thus, action planning (often) is an off-line process that, as **Wolters & Raffone** rightly emphasize, requires some kind of short-term memory capacity. As sensorimotor processing does not stop while planning is underway (e.g., planning an utterance does not require to stop walking), action planning seems to occupy an extra input-output loop; a loop that can be temporarily decoupled from on-line processing and reconnected to the processing stream whenever necessary (a strategical advantage emphasized by **Bryson**). This suggests an architecture such as sketched in Figure R1, where

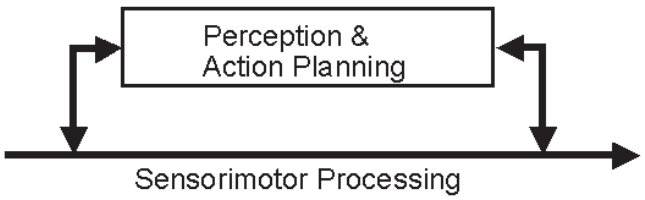



Figure R1. Sensorimotor processing and action planning take place in parallel. The on-line stream of sensorimotor processing is constrained by perceptions and action plans that are elaborated in parallel (though not necessarily synchronously).



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

an on-line stream of sensorimotor processing proper is constrained by perceptions and action plans worked out in a parallel, but not necessarily synchronous perception-action stream (we have to admit that our target article is less clear about this aspect than it should have been). This picture is not too different from that drawn by **Pisella et al.** Moreover, aspects of the former stream may be likened to Milner and Goodale's (1995) dorsal action stream, aspects of the latter to their ventral perception stream (see **Westwood & Goodale**).

Second, as action planning subserves a particular goal it can be expected to determine the major, goal-relevant features of a given action but it does not need to, and in many cases will not, specify all features of the (sub-)movements making up the action (Jeannerod 1984; Turvey 1977). Take, for example, the studies of Goodale et al. (1986) and Prablanc and Pélisson (1990), who asked participants to move their hand from a starting position to a goal location indicated by a small light. Unnoticed by the participants, in some trials the light made jumps of several degrees at the beginning of the movement. Nevertheless, the movements were carried out correctly and without any observable changes in their timing. As re-planning the movements should have produced some measurable time costs, this suggests that the changing movement parameters were not planned in advance and then re-planned but, rather, were specified by continuously fed-in visual information that steered the hand towards the target. To do so, however, relevant stimulus features must have been specified and linked to action parameters so as to delegate the movement's fine tuning to lower-level sensorimotor channels. Similar signs of control delegation are reported by **Pisella et al**. They show that moving a target stimulus can evoke fast adjustments of the participant's hand position even when target motion actually signaled stopping the movement. Although alternative interpretations are possible (e.g., participants may have forgotten or confused instructions, De Jong 2000; or have started to move before issuing the stop signal to the motor system, see Logan 1984), such an observation fits nicely with those of Goodale et al. (1986) and Prablanc and Pélisson (1990). However, why **Pisella et al.** feel that "automatic corrective movements contradict Hommel et al.'s claim that stimulus presentation is necessary but insufficient for response to occur" remains unclear to us. The sentence in our target paper that follows the one Pisella et al. refer to says, rather clearly, that "nothing will happen upon stimulus presentation until the participant has been instructed to respond in a particular way and he or she is willing to do so." Accordingly, we predict that Pisella et al.'s participants would not have made corrective movements had they not been instructed to carry out pointing movements toward the presented targets at all. If this not overly risky prediction holds, the corrective movements can be seen as a consequence and, in fact, as an indicator of the participants' action goals, which is in full agreement with TEC.

Third, planning an action requires knowledge about how a given goal can be achieved (Elsner & Hommel 2001; Hommel 1998a). Depending on the particular action and the environmental conditions, such knowledge may not be necessary for the mentioned fine-tuning via sensorimotor channels. Indeed, and here we agree with ecological approaches, there may be sufficient information "out there" to specify the parameters of a long jumper's final few strides to the take-off board (**Kim & Effken**; see also sect. R3.4). Yet, the actual and efficient use of this information depends on a whole number of preparatory processes: the very idea of jumping into that particular sand-pit, to do so by running towards it, to jump off with the dominant leg, to lean the body forward while in flight, and so forth. Most of these preparations are likely to be carried out way ahead of time of contact with the take-off board, so that anticipations of later parts of the action can shape earlier parts to make the whole action more efficient (**Rosenbaum**; Rosenbaum et al. 1990). Thus, they take place before the environmental or body-related information they refer to is available, which means that they must depend on some kind of internal knowledge – be it an internal model or stored instances of movements (**Rosenbaum**), or of movement-effect episodes (Hommel 1997; 1998a; **Kunde**).

To summarize, we distinguish between an on-line sensorimotor stream of information flow, that TEC does not cover, and a (potentially off-line) perception-action stream, that TEC does cover. We assume that while the former is affected and constrained by action planning, the planning processes themselves take place in the latter. Even if action planning may often be selective in not specifying kinematic peculiarities that can be more precisely specified by on-line environmental information, we would not feel comfortable with **Westwood & Goodale**'s distinction between "deciding *what* to do" (a job they seem to ascribe to Milner & Goodale's [1995] ventral pathway) and "specifying *how* to accomplish an intended action" (ascribed to the dorsal pathway). Let us take their example of picking up a cup of coffee. We can see that deciding to perform that action at all is a What-decision. But what about choosing the hand, the fingers performing the grip, the speed towards the cup, the force of the grip, the part of the cup that is grasped, the speed with which the cup is brought to the mouth – are these all What-decisions as well? If not, are they all *exclusively* performed via the dorsal pathway? Would that not imply that none of these decisions is open to voluntary, or at least not conscious, control? In our view, such a distinction raises more questions than it answers, and it becomes even worse if we consider the rather common situation that the cup is lifted without being looked at.

### *R2.3. Codes*

TEC aims at describing relevant functional characteristics of the codes underlying perceptual and action-planning processes. In speaking of codes representing perceived and to-be-produced events, we make what is in our view a rather uncontroversial assumption: that perceiving an external event is associated with a correlated change in the perceiver's cognitive state (e.g., as indicated by neural activity) and that producing an external event is preceded by a correlated change in the actor's state. There are many ways to describe such changes and some levels of description may be better suited for particular purposes than others. Apart from the functional, analytic level of description we preferred in presenting TEC, one may focus on the activation levels of neuroanatomical structures (**Chaminade & Decety**; **Cisek & Kalaska**) or model neurons (**Wolters & Raffone**), interaction patterns between or within neural assemblies (**Chown et al.**), or characteristics and changes of cortical maps (**Dinse**), and one may even wish to consider concentration of neurotransmitters.


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

Although it is an all but trivial task to properly relate these different descriptions of representational codes to each other, many commentators demonstrate that it is feasible and, indeed, the whole idea underlying cognitive neuroscience strongly depends on this. Importantly, whatever the description level chosen, we do not regard neural codes to be properly characterized as "mental" (**Richardson & Michaels**) – a term that, apart from citations, we did not use at all – or "nondynamic" and "time-free" entities "that seem to sit outside of natural law" (**Shaw & Wagman**), and we see no reason why TEC would motivate such characterizations. Moreover, we do not see the codes assumed by TEC to be properly characterized as probabilistic anticipations on the basis of which people form "an expectation of what might occur" (**Kim & Effken**). Instead, we fully share **Millikan**'s view that the codes underlying perception and action planning have a double-faced function in both representing a particular state of affairs *and* telling what could be done about it. Hence, in Millikan's own words, "the same complete representation token can have two functions at once, being both a fact-presenter and an action-director." The only thing we need to add from a TEC point of view is that, given the assumption of distributed representations, it may still be possible to analytically or even physically (e.g., by means of single-cell recordings or through lesions) decompose those tokens into smaller units, that is, feature codes.

This assumed high degree of integration of perceptually derived and action-related codes distinguishes TEC from stage models with their implicit or explicit separation of codes. Therefore, we doubt that the mere observation that both types of models use the same term (**Proctor & Vu**) signals any deeper theoretical connection; in our view, this merely reflects that they both presuppose (as any approach in the wider field of cognitive neuroscience does) that there is some internal activity correlated with external events. With respect to the assumed integration of codes across perception and action **Pisella et al.** are concerned that TEC may not be able to account "for the fact that the same perceptual event can give rise to several actions (e.g., designating, grasping, squeezing)" and that "the perception of a distal stimulus" might "imply that one has already selected a given action to perform on or with this object." In our view, these commentators overlook major assumptions we have made. Most importantly, TEC's prototypical agent does not passively await some stimulus information and then makes a decision how to react thereupon. Outside psychological laboratories people commonly look out for particular stimulus events they are interested in, which means that perception is as intentional as action is (**Jordan**; see sect. 2.1). If so, the actual processing problem is exactly contrary to that posed by **Pisella et al.**; it consists of finding a stimulus that matches the current interests and action goals, not in checking out what action a given stimulus may afford. Hence, when looking for something to write, a pen simply matches one's current goals, and for neurophysiologically healthy adults there is no "necessity" to bother about, such as to suppress squeezing or throwing the pen. How this kind of goal-directed coding of perceptual events may be achieved has been discussed in section 3.2.3 of our target article.

Another concern raised by **Oriet et al.** (and, to some degree, by **Proctor & Vu** and **Sanders**) relates to the question: how original is our assumption of common codes? In particular, **Oriet et al.** ask "whether a common coding model can be distinguished from a classical model in which interactions between perception and action codes are postulated." Apart from the terminological confusion we discuss in R3.2, our first answer to this is: If one really postulates multi-lateral interactions between perception and action codes one would no longer defend a classical model, as it is this very uni-directional flow of information from input to output that, in our view, characterizes classical models (see **Proctor & Vu**). But, more concretely, assume a hypothetical, purely (classical) perceptual code P and a purely action-related code A. Let them represent, for a simplified example, the fact LEFT, so that P would become activated if a "left" stimulus is perceived and A would be activated if a "left" action is planned. Were they independent, it would be difficult to understand why left actions are primed by left stimuli, and vice versa, to name just one example from our empirical review. To account for such mutual priming one at least needs to connect P and A by some kind of association, which seems to be what **Oriet et al.** suggest. However, how would such an association account for the observation that planning a left action impairs the perception of left stimulus (Müsseler & Hommel 1997a; 1997b) – an effect that Oriet et al. were able to replicate and extend? That is, why should using A *interfere* with using P? To account for that, one would need to link P and A to a degree that gets at least very close to assuming some kind of functional unity. Once this point is reached, we suspect there is not much left for major theoretical arguments.

A final question with regard to the relationship between the codes used for perceptual events and for action plans is raised by **Sanders**, who asks whether "perception and action planning also share a common representation in a Sternberg classification task." The answer is as simple as it is general: That depends entirely on what the stimuli and the responses are. As perceptual events and action plans are coded in terms of distally defined features, code overlap (or partial code identity) exists to the degree that the distally defined features of the stimulus and response in question are the same.

### *R2.4. Features*

TEC assumes that perceived and to-be-produced events are coded in terms of their distal features, that is, represented through activated and integrated feature codes. We took pains to point out that TEC allows for the coding of any feature, be it as "simple" as the pitch of a sine tone or as complex as a chair's "sit-on-ableness" – as long as it can be discriminated in perception and/or action planning. The reason for so liberally considering even "arbitrarily chosen, arbitrarily combined features," as criticized by **Galantucci et al.,** has to do with real life. In contrast to the seemingly naturalistic picture drawn in most ecologically inspired commentaries, mastering real life in modern Western cultures involves a multitude of tasks with arbitrary combinations of perceptual and action-related features. This includes, for example, pushing the right button to quiet the alarm clock in the morning, brewing coffee by using rather complex electronic equipment, and driving a car while navigating through a whole world of traffic lights and linguistic instructions. Most people manage to master these tasks, and we want to understand how they can. As the stimuli



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

they face are often arbitrary combinations of simple and complex features of perceivable events, and as the actions they perform are often arbitrary combinations of movement parameters and other movement elements, we think it is important to build a theory that helps understanding how people code both natural *and* artificial events. We definitely agree with ecological approaches that the search for effectively used information in perceptual events and produced actions is not easy, and we also agree that higher-order variables like tau (**Kim & Effken**) or even more complex derivates from motor competencies (**Galantucci et al.**) may be involved in specifying some parameters in jumping movements and speech perception, respectively. Yet, we doubt that accounting for the most of our everyday perceptions and actions will be possible without considering less "natural" stimuli and actions than those favored by ecological psychologists. This is why TEC allows for both arbitrary and "natural" feature combinations in perception and action planning.

A drawback of this liberal stance is that we are unable to provide *a priori* definitions or predictions of which features are coded under what circumstances (**Hochberg**; **Wolters & Raffone**). This is not so much a problem in experimenting, as participants can be instructed and tasks tailored to highlight particular features and make them task-relevant. However, outside the lab such tight control is not possible so that predictions necessarily lose precision. Moreover, it may be that the reliance on one or another feature depends on and, thus, varies with practice and expertise (**Sanders**) – just think of processing the taste of wines or the properties of snow. This is not a problem unique to TEC. What features people attend to, select, and eventually learn do always depend on a mixture of relatively easy to objectify task constraints and contextual circumstances, and much more difficult to determine individual factors having to do with abilities, skill level, learning history, attentiveness, and so forth. Hence, there will always be factors whose effects are easier to predict than of others, a problem TEC shares with any other model in perception, categorization, memory, or motor programming. Considering this, we are sceptical with regard to the possibility of identifying and codifying feature codes in an *a priori* fashion, as demanded by Hochberg or Wolters & Raffone. Instead, we prefer to stick to a circular definition: Feature codes can code anything that a perceiver-actor is able to discriminate in perceiving and/or producing an event. As our empirical review and the new evidence in the commentaries demonstrates, this does not necessarily prevent one from making successful predictions.

Another assumption we make is that feature codes refer to distally defined information, not to the sensory channels this information has been received by. **Vogt & Hecht** have taken this to mean that TEC only allows for coding "abstract" information and therefore is unable to account for surplus information if different channels are involved. This is true *only if* all sensory channels would deliver exactly the same type of information in exactly the same quality. As this is not a reasonable assumption – just consider localization by eye versus ear, or texture identification by eye versus hand – TEC is well equipped to deal with findings showing both transfer between, *and* various contributions from, different modalities. In a way, sensory channels are not too different from TV channels: the information they deliver points to external facts, not to the channels, and multiple channels increase the amount of information one gets. Along the same lines, TEC does not really introduce the problem of differentiating between seen and performed action, an issue raised by **Chaminade & Decety**. As long as the perceiver is not the actor it is very unlikely that the information about a seen and a performed action is identical; just think of the rather specific information delivered by kinesthetic channels, or the way action goals (which are absent when perceiving an action) "color" perception and action coding through feature weighting.

An interesting, additional issue with respect to feature codes is raised by **Meiran** and **Richardson & Spivey**. In our target article we focus on perceptual events and action plans, in other words, on events outside the body. But what about internal codes referring to intended (but not yet achieved) events, (perceived) emotions, semantic, linguistic, and other memory contents? Should TEC allow for such codes to become integrated into event codes, thereby coloring, so to speak, the coded perceived event or action plan? Albeit very briefly, we did signal this possibility in the target article (sect. 3.2.1) but did not develop this issue in the empirical review – simply because at that time we were unable to find data speaking to it. However, there are several recent observations that encourage us to more strongly consider the integration of such "internal," task-specifically weighted feature codes. For one, there is **Meiran**'s own recent work, in which he successfully applies his idea that stimuli are integrated with the context-dependent meaning of their response, to a number of findings in the area of task-switching (e.g., Meiran 2000b; Meiran et al. 2000; see also Hommel et al. 2000b). Indeed, the possibility of binding codes of already known stimulus and action features before an action takes place would provide ideal support for prospective memory: Once prepared, such a binding would be able to control the more or less automatic execution of an action under appropriate circumstances (Bargh & Gollwitzer 1994). Evidence for the integration of emotional information is also emerging. For instance, actions that are consistently followed by a mild electric shock (Beckers & De Houwer 2000) or a "grumpy" face (Van der Goten et al.; Hommel 2001) have been demonstrated to become increasingly compatible with word stimuli of negative emotional valence, and comparable effects occur for positive action effects. This suggests that actions become integrated with codes of positive or negative outcomes, which then play a role in stimulus-driven response selection (action induction). Interestingly, this fits well with Damasio's (1994) concept of a "somatic marker." Finally, as reported by **Richardson & Spivey**, even semantic and linguistic information has been found to become integrated with perceived events and action plans. Taken together, the evidence strongly supports our broad interpretation of what counts as an event feature.

### *R2.5. Events*

In principle, TEC is intended to account for coding events of any sort and on any time scale. Yet, most examples in our empirical review refer to rather primitive stimulus signals and relatively simple laboratory actions – although we think that the range of actions is actually wider than some commentators admit and that, in contradiction to **Pisella et al.**'s implicit assumption that only pointing and grasping actions have goals, they were all goal-directed. Still, many perceived events and performed actions in our daily life are


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

richer and more complex, which raises the question of whether we can really extrapolate the available evidence to larger-scale events. There are two related issues that become increasingly important when turning to such events: (1) they often possess some kind of (at least perceived) hierarchical structure (**Zacks**), such as a scene consisting of parts and sub-parts, or a multi-layered action like preparing for an exam; and (2) that they typically comprise a number of sequential steps (**Chown et al.**), such as in making coffee or in watching a movie. A hierarchical and sequential structure introduces coding problems that go beyond TEC's present capabilities. Very likely, coding complex events involves some kind of schematizing (**Zacks**) or chunking (**Chown et al.**; **Lane et al.**; **Wolters & Raffone**) of their parts, and we agree with the commentators that TEC would (and should) benefit a lot from relating to available accounts of hierarchical event coding, such as PLAN, CHREST, and Zacks and Tversky's (2001) model. But TEC also has something to contribute to understanding the coding of multi-layered events. For instance, TEC-driven studies have shown that the cognitive representations of elements of complex visual arrays are determined by whether, and how, their perceptual and action-related features overlap. In particular, houses of map-like configurations are integrated into the same cognitive cluster if they share shape- or color-related features (Gehrke & Hommel 1998; Hommel et al. 2000a) or the action they signal (Hommel & Knuf 2000; Hommel et al. 2001).

With respect to the coding of sequences, we see no difficulty in assuming that elements of a sequential action are temporarily linked and organized by means of syntactic structures, such as the action-sequencing cells discussed by **Bryson**. As the order of elements is a perceivable as well as plannable feature of events, we find it only natural that neural means exist to code these features. **Hartsuiker & Pickering** even report evidence that syntactic features obey the same processing rules and, thus, give rise to the same types of phenomena than the more content-related spatial and figurative features covered by our review. Other challenges may be more difficult to meet. For instance, if action sequences get longer only the first few elements tend to be fully prepared before the sequence is started, whereas later elements are planned while execution of the preceding elements is underway (e.g., Semjen & García-Colera 1986; Van Donkelaar & Franks 1991). In TEC terms this would imply a mixture of fully integrated action elements and elements that are only activated (or maintained), and we do not see how TEC could predict how many elements fall into each category or how such a composition would behave. Yet, irrespective of such open questions, we have seen no convincing argument that a TEC-inspired approach is *in principle* unable to deal with more complex actions. Quite to the contrary, the work on the processing of sentences and stories reviewed by **Zacks** suggests that TEC fares reasonably well if applied to events on larger time scales.

### *R2.6. Activation and integration*

TEC distinguishes two basic representational states a feature code can take on, activation and integration. Facing a perceptual event or action opportunity one is currently not interested in leads to a brief and quickly decaying activation of the codes corresponding to this event's features (see Figure R2, feature code $f_3$ in Phase II). Interest in the event is associated with an (attentional-intentional) increase in the weights of its interest- or goal-related features (the feature-weighting principle), which again increases the net activation these codes reach when coding the event in question (see $f_1$ and $f_2$ in Phase II). This additional boost increases the likelihood of codes reaching an integration threshold that determines which codes become part of the integrated event code (see $f_1$ and $f_2$ in Phase III). If the binding dissolves, the activation of the previously bound codes will start decaying (see $f_1$ and $f_2$ in Phase IV) and eventually go back to baseline.

A whole number of commentaries addressed the several problems and virtues of these process-related assumptions. We have a twofold reply to their concerns. On the one hand, we are fully aware that TEC is still too underspecified to deal with the details of many empirical phenomena, and in several cases is simply too powerful (i.e., nonspecific) to allow for clear-cut predictions and rigorous experimental testing. In particular, this applies to temporal assumptions regarding when the hypothetical phases take place and their durations, as well as to the possible preconditions for these processes and for aspects of their temporal behaviour (**Diedrichsen & Hazeltine**; **Kunde**; **Oriet et al.**; **Zacks**). This is certainly a weakness we have to admit, but we think we can (and, for the moment, should) live with it. For one thing, these missing assumptions do not really touch TEC's basic architecture or the logic of its operations, and we therefore prefer to resolve the questions relating to them empirically. Indeed, it is not unlikely that as far as experimental tasks are concerned, the necessary temporal parameters are strongly dependent on the particularities of the task, the stimulus material, the type of responses, and the strategies of the participant. For instance, the time when integration starts and how long it takes may vary with the number of stimuli competing for integration, the complexity of a planned action, or the time-scale of the event in question

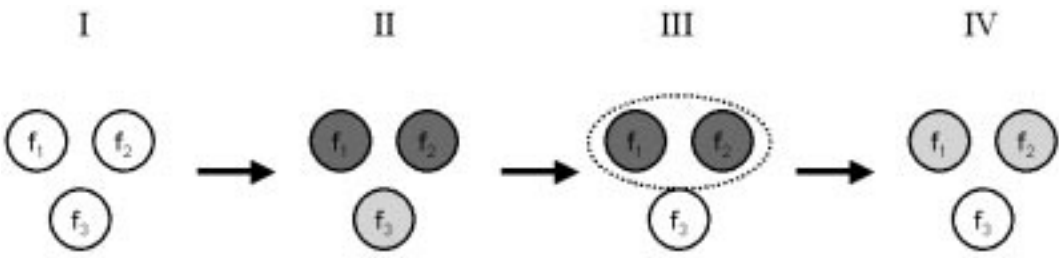



Figure R2. Four sequential phases of feature activation and integration where one feature ($f_3$) code rapidly decays after activation (Phase I) whereas activation for $f_1$ and $f_2$ increases (Phase II), yielding an integrated event code (Phase III) and starts to decay again (Phase IV).



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

(**Zacks**). How long it takes for a previously bound code to decay will depend on whether or not the perceiver-actor anticipates making use of this code in the near future. And people may be able to adjust, and therefore vary in, the size of their temporal "integration window," that is, the interval across which evidence is sampled into the same event file (Lupiáñez & Milliken 1999). We do not see how all these and other possible complications can be resolved on the basis of *a priori* theoretical principles, and therefore do not find it useful to unnecessarily restrict TEC's bandwidth by including some arbitrarily chosen specifications on the basis of preliminary evidence. Rather, we see TEC as a framework to motivate and guide the construction of a whole number of more task- or domain-specific TEC-type models that do lay down such specifications and can, therefore, be empirically tested with more rigor. We even consider it useful to construct competing TEC-type models, that is, models that share their basic architecture but differ with respect to their parameters.

On the other hand, the hypothesized picture of an interplay of activation and integration has recently found considerable support in quite a number of tasks and areas other than those considered in our review, spanning the integration of eye movements and linguistic material (**Richardson & Spivey**), dual-tasks combining eye and hand movements (Stuyven et al. 1999), judgments of object properties (Richardson et al. 2001; Tucker & Ellis 1998), and the integration of lexical and syntactic information (**Hartsuiker & Pickering**). Moreover, even the present version of TEC is not as underspecified as some commentators claim; there were quite a number of challenges we found not too difficult to deal with even in the absence of more task-specific versions.

For instance, although we consider it possible that additional assumptions and qualifications may be required to apply TEC to the behaviour of mentally ill people, we do not see how the observations of Müsseler and Hommel (1997a; 1997b) can be explained by conditioned blocking (**Oades & Kreul**). Blocking requires some amount of training, hence, a number of repetitions of particular stimuli and/or responses; yet, Müsseler and Hommel's action-effect blindness is a trial-to-trial phenomenon that appears in experiments with constantly changing combinations of stimuli and responses.

We also are unable to see why, and in what sense, the assumption of distal coding conflicts with the observation that action planning is affected by the relation between the hand's starting position (a well-defined distal event) and (other) stimuli (**Pisella et al.**). With respect to action planning TEC uses the terms "proximal" and "distal" to refer to specifications of a planned event in terms of muscle innervations and to-be-perceived, external attributes, respectively; this is not to be confused with "close to the body" versus "far from the body." Moreover, it is true that TEC has no full-fledged theory of attentional selection built in, which makes principled predictions about the coding of distractor stimuli difficult; yet, the observation that trajectories of hand movements are affected by concurrently coded stimuli (Pisella et al.) seems to be so much in line with TEC that we fail to see a challenge here (see also sect. 4.3.2 in the target article for how TEC accounts for effects that seem to indicate inhibition).

**Oriet et al.** report a number recent studies from their lab that replicate and considerably extend the observations of Müsseler and Hommel (1997a; 1997b). However, in contrast to Stoet and Hommel (1999) and Wühr and Müsseler (2001), they found reliable interference between a planned action and a masked stimulus not only before, or while, but even after the action was carried out. As they say, this is inconsistent with our assumption that elements belonging to an action plan become automatically desintegrated immediately after use. Instead, participants may have kept their plans active for some longer time than in our own studies. Yet, even in the studies mentioned by **Oriet et al.** dramatic reductions of the action-induced impairment of perceptual identification were observed as soon as the action was fully executed, and in some conditions the effect was indeed eliminated. Hence, the empirical differences are not as drastic as **Oriet et al.**'s commentary might suggest and it is not too difficult to explain them. Moreover, even though we agree with Oriet et al.'s general argument that more specification is needed for a more task-specific modeling, we think that they overestimate the degrees of freedom TEC provides for predicting and accounting for empirical findings. Activation and integration are states that are strictly bound to particular functions, and this in principle allows for independent testing of even post-hoc assumptions. For instance, if the observation of blindness effects after execution of an action really results from participants maintaining their action plans somewhat longer, one should find an increase or decrease of this pattern if one rewards for or punishes the maintenance of those action plans, respectively (e.g., by introducing a tertiary task in-between action execution and stimulus-related judgment calling for frequent repetitions versus alternations of the previously planned action). Moreover, if the plans are really maintained longer, one might find effects on the speed of responses carried out thereafter, for example, in the stimulus-judgment part of the task – especially if this part follows shortly after and especially in participants showing stronger indications of "maintenance." In a nutshell, TEC does not in fact "as easily account for a pattern of results as it can for the exact opposite pattern" (Oriet et al.) because the additional assumption one needs to make to account for an unexpected outcome is open to an independent empirical test.

**Kunde** also points to some, in his view, inconsistencies between TEC's predictions and its empirical observations. In particular, he discusses Craighero et al.'s (1999) finding that planning a grasping movement primes the processing of a "Go" stimulus that looks like the grasp object, and he asks whether TEC should not have predicted negative effects here. However, there were some major differences between Craighero et al.'s and Müsseler and Hommel's tasks. People in the former prepared one action directed towards an object, that is, a single sensorimotor event to be triggered by the "Go" stimulus. Here, the processing problem consisted in maintaining and then launching a single representation of an object-oriented movement, but no separate stimulus-related response was required. And action-congruent "Go" signals did not just feature-overlap with the goal object – they perfectly matched it, so that no coding conflict could arise. In contrast, the Müsseler and Hommel task requires the formation and maintenance of two different, independent cognitive structures, an action plan and a stimulus representation for later report. The processing problem in this case is therefore not restricted to one structure and mere maintenance, but includes keeping the two maintained structures separate. Also, the "object"


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

of the planned action was a response key, which feature-overlapped but was otherwise very different from the to-be-identified stimulus. Hence, in contrast to Craighero et al.'s task, where people might have used the same event code for coding both the stimulus and the planned action, the major problem for Müsseler and Hommel's participants consisted in creating one event code while maintaining another, and in keeping them apart although they are linked via a shared feature code. Accordingly, we do not consider the findings of Craighero et al. a real challenge to TEC. The same applies to Hommel and Schneider's (in press) observation that selecting a manual response primes the selection of a bar-marked element of a small search array. As reported in that study, there was strong evidence that the actual stimulus selection took place after response selection and execution was completed. Thus, under the assumption that plan elements were quickly unbound (Stoet & Hommel 1999), one would expect that the (decaying) activation of plan elements could bias the eventual stimulus selection. And this is what happened. The problem here is not that predictions from TEC would be ambiguous, it is just difficult to determine *a priori* precisely *when* the assumed processes take place if a task gets complicated.

The objections of **Diedrichsen & Hazeltine** are similar to **Kunde**'s, and so is our reply. We have already pointed out how Hommel and Schneider's (in press) findings fit into the picture. Our interpretation of Diedrichsen et al.'s (2000) observations is somewhat different. What they found is that the distractors that are compatible or incompatible with a target (and the implied response) have a stronger impact if they appear on the side where the correct response is carried out. In our view, during the first, activation phase target information is continuously fed into activation of the associated action-feature codes (Hommel 1998c; Hommel & Eglau, in press), which include response location. Activating location codes (access to which is shared by stimuli and responses) "backward-primes" the stimuli sharing the activated code, so that response-compatible distractors receive a processing advantage. We found a similar phenomenon under dual-task conditions: Hommel (1998c) observed that compatibility between the response of the secondary task (e.g., the verbal response "red") backward-primes the stimulus of the first task (e.g., a red-colored stimulus). Thus, these effects are not restricted to the spatial domain. Note, however, that all these priming-type effects are observed some time *before* the response in question is carried out. If we assume that in these reaction-time experiments response execution immediately follows response planning, this means that priming is observed in the earlier planning phase, that is, the activation phase. Accordingly, TEC would predict that the effects should differ from a situation in which people have several seconds to plan their response before facing the to-be-processed stimulus, as in Müsseler and Hommel's (1997a; 1997b) studies. And this is what Diedrichsen et al. (2000) observed.

We found **Diedrichsen & Hazeltine**'s limited success in applying TEC to the concurrent or temporally overlapping planning of multiple actions, more challenging. On the one hand, the findings of Stoet and Hommel (1999) can be replicated with combinations of eye and hand movements (Stuyven et al. 1999), which rules out **Kunde**'s objection that body instability may have played a role and demonstrates some degree of generality. On the other hand, however, the discussion of **Diedrichsen & Hazeltine** reveals the (admitted) difficulty in defining what an event is. The two actions planned in Stoet and Hommel's task were separated by several seconds, and therefore clearly required the creation of two different plans. However, with decreasing temporal separation it becomes unclear whether people still create two plans or somehow merge these into one coherent structure. Even if one introduces stimulus-onset asynchronies (SOAs), the often short interval between the actions may still motivate people to either use one coherent plan, or re-use the previous plan by only modifying the changed parameters (Rosenbaum et al. 1986). Then predictions from TEC become muddy, the more so as it does not provide the means to predict which strategy is used under which conditions. Which is one more reason to point out that we see TEC only as a guide to build task-specific models, not as a substitute for such models.

Finally, **Wolters & Raffone** have some objections to our distinction of activation and integration processes, and they discuss reasons why and how these two processes might interact. Indeed, it makes sense to assume that integrating a feature code into a coherent event code impacts upon its activation level and thereby prolongs its "lifetime." Conversely, it seems obvious that only (sufficiently) activated codes can become integrated. Yet, this does not, for one, rule out the possibility that activation and integration phases have different effects on other codes, and we think that the demonstration of both positive and negative effects of feature overlap supports our assumption that they do. Moreover, **Wolters & Raffone** argue that

> the selective allocation of a shared feature . . . to only one of two concurrently activated event representations, cannot be plausibly based on the oscillatory *synchronization within* and *desynchronization between* the two event representations: since synchronization implies a transitive relationship, shared nodes may lead to interference-prone coding and readout of the events.

However, apart from the fact that TEC is not tied to oscillatory synchronization as a physiological implementation of integration, the presence of interference between action planning and perceptual identification reported by Müsseler and Hommel (1997a; 1997b) in fact points to some kind of "interference-prone coding and readout of the events." Hence, rather than challenging TEC's assumptions **Wolters & Raffone**'s commentary in our view provides additional support for the logic underlying them.

### *R2.7. Control*

TEC focuses on *how* events are cognitively coded but it doesn't have much to say about *what* events are considered under particular circumstances. Thus, it does not (yet) include an elaborated theory of selection, so it cannot satisfactorily deal with issues like the selection of targets from distractors (**Pisella et al.**), or the selection among competing goal-satisfying behavioral alternatives (**Olivetti Belardinelli & Basso**; **Bryson**; **Kunde**). However, it is also true that even the present version of TEC does have some control built in. In particular, we assume that goal-related features of objects and action plans are weighted more strongly than other features. This not only contextualizes the emerging representations of events, it also affects their impact on behavior by granting goal-relevant objects and action plans a processing advantage. As discussed by **Meiran**, feature weighting may suffice to explain great portions of inten-



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

tional task-switching – an ability commonly considered to reflect intentional control – and influential cognitive-control models such as that of Cohen and colleagues (Cohen et al. 1998; Cohen et al. 1990) operate basically on the same principle (Hommel et al. 2000b). With respect to input selection, the assumption that goal-related events receive selective top-down support which again biases the competition for selection, fits well with recent models of attentional selection, as those of Bundesen (1990) or Duncan and Humphreys (1989; Duncan 1993), and **Ivanoff & Klein** rightly point out further similarities with Norman's (1969) and Folk et al.'s (1992) ideas on attentional control. True, these connections need to be worked out in more detail in the future; but given that TEC was not meant to be a control model proper, it does not compare too badly with those that are.

#### *R2.8. Neurophysiological and anatomical basis*

TEC is formulated in purely functional terms with no particular connection to the brain hardware through which the hypothesized functions may be implemented – a decision regretted by some commentators (e.g., **Chaminade & Decety**; **Pisella et al.**). There are three reasons that in our view justify this decision *at this point.*

First, although some, certainly encouraging, neuroscientific studies on perception-action relationships are available already (and some were indeed considered in our target article), this area of research is still in its infancy. For instance, most of the relevant neuroscientific studies **Chaminade & Decety** and **Pisella et al.** discuss either just appeared or are still in press, and the validity of some central findings underlying Milner and Goodale's (1995) distinction between a perception and an action stream is under heated debate (e.g., Bruno 2001; Carey 2001; Rossetti & Pisella 2002). So, before trying to map psychological function to biological substrate it seems safe to await some degree of consolidation in findings and interpretations on the biological side.

Second, as with any research, what neuroscientific approaches can find is constrained by the methods they employ. The machinery and techniques dominating the current discussion of perception-action relationships (as well as the arguments of **Bryson**, **Chaminade & Decety**, **Cisek & Kalaska**, **Dinse**, and **Pisella et al.**) focus on the activation of either single cells, or cell assemblies, or whole cortical areas; or, as in patient or lesion studies, on the presence or absence of those areas. It makes sense to assume that the activation of those neuroanatomically defined units speaks to the functional process of activation postulated in TEC. However, we are not so sure whether they, in principle, can tell us something interesting about communication between units and integration of the information they deal with. This may work if integration is achieved by convergence on single grandmother cells or grandmother assemblies, which might be localized in a particular area or system, which can then be detected via brain imaging, lesions in this area, or single-cell recordings. Yet, if integration is achieved by coordinating the behavior of neurons or neuronal assemblies, as **Wolters & Raffone** or Singer (1994) suggest, it need not lead to any detectable increase, perhaps not even to a change, in brain activation (as timing and average firing rate are logically independent). If so, it may turn out to be extremely difficult, if not impossible, to "localize" integration processes in the brain, and conclusions based on techniques that focus on localization may have limited value. Therefore, it seems safe to wait and see how findings from neuroscientific "activation" analyses fit with results from techniques better suited to reveal communication processes, such as mass-cell recordings or magnetoencephalography.

Third, given these and other considerations, we as yet do not see the promise of over-additive insights when mapping functional process descriptions to brain structures. Of course, it is highly interesting to see that the general principles claimed in TEC also hold for, and prove to be useful to understand, the neurophysiological underpinnings of perception-action relationships. We are therefore grateful to the commentators for having pointed out these parallels and possible points of connection and correspondence between function and substrate; the review by **Chaminade & Decety,** especially, is very encouraging in this respect. We also have no reasons to object to the particular ways suggested by these commentators to map function to substrate, and to direct our attention to probable TEC-related systems in the following regions of the brain: in the parietal lobe, where information about stimulus-response relations may be stored (Chaminade & Decety, **Pisella et al.**); the prefrontal cortex, which may be involved in "contextualizing" and biasing events (**Cisek & Kalaska**; Cohen et al. 1998); the cerebellum, which may mediate perception-action integration (Pisella et al.) and forward-modeling of action (Cisek & Kalaska); and the basal ganglia, which may contribute to the selection of action plans (**Bryson**). Nevertheless, what would happen if these suggestions prove incorrect? What if, to take a fictitious example, we explicitly assumed that goal-related feature weighting is "performed by" the prefrontal cortex, and then be faced with indications that features are still weighted in patients or animals after complete removal of this cortical structure? Should we then give up the assumption of feature weighting or the assumed connection between weighting and prefrontal cortex? We believe that most people will find the latter more reasonable, simply because understanding a psychological process does not critically depend on the cortical location where the process takes place – if it can be localized at all.

These arguments are by no means intended to encourage mutual ignorance between psychological and biological approaches. On the contrary, we find the developing neuroscience of perception and action (for a label) extremely interesting and the emerging parallels to functional models very inspiring. Nor do we exclude more fruitful integration of functional and physiological ideas in the near future. It is just that we, at this point, do not yet see sufficient justification for (nor the synergetic effects of) tying TEC to particular cortical structures.

### R3. Relations to other approaches

#### *R3.1. Historical predecessors*

In our target article, on several occasions we pointed out that TEC does not come out of the blue but, on the contrary, takes up a number of elements from other, in some cases surprisingly old, theoretical approaches. Yet, we may


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

have failed in making some important connections sufficiently obvious (**Hochberg**), such as the links with the Tolman theory of purposive behavior (discussed in more detail in Hommel 1998a) and Gibson's ecological approach (see sect. R3.4). **Sanders** mentions further interesting connections to the work of von Weizsäcker and Werner, and there are some more points of contact to other ideas that the interested reader will find elsewhere (Aschersleben & Prinz 1995; Elsner & Hommel 2001; Hommel 1997; 1998a; Müsseler 1999; Prinz 1984; 1987; 1992).

TEC's connection to Tolman (1932; 1959), also emphasized in the commentary of **Chown et al.**, is instructive with respect to three issues. First, apart from the distal coding issue discussed by **Hochberg**, TEC owes to Tolman the insight that feedback (i.e., action effect) does not only possess a motivational function (by rewarding or punishing for the action) but also provides *information* about what the action leads to (Tolman et al. 1932). Accordingly, it makes sense to assume that action effects do not only supply the glue for linking stimulus to response (Walker 1969), but that their representations become integrated with the response pattern itself. This way the effect code informs about means-end relations and thereby serves as both a forward model (**Cisek & Kalaska**; **Rosenbaum**) and a retrieval cue of the action (Harless 1861; James 1890; Lotze 1852).

Second, the major historical problem of Tolman's approach, at least in the area of learning theory, is that it nicely accounts for what rats (and, by inference, humans) may know about events and environmental conditions, but it fails to explain how and why it ever starts moving. Indeed, if one considers the information a rat or human acquires by observing another as pure "cognition," that is, as mere information laid down in some storage system, one needs to introduce additional processes that interpret and make use of this information to carry out some overt behavior. That is, one needs some machinery that translates perceptual codes into action patterns, as is postulated by classic stage theories of information processing (see sect. R3.2). This is certainly feasible, but then one faces all the problems translation approaches entail (as discussed in our target article), like explaining why some translations are easier than others, and why action planning affects perception. But such problems do not arise if one considers codes to have a direct, active impact on action control (**Millikan**), as TEC suggests.

Third, an important driving force for Tolman's cognitive approach was the difficulty to explain maze learning in animals on the basis of direct stimulus-response associations. For instance, if rats can transfer maze-specific knowledge acquired by wading to swimming, or vice versa (Macfarlane 1930), they do not seem to have learned particular stimulus-response chains. Instead, what they acquired must have included some kind of representation of the locations visited, or, as TEC suggests, a sequence of perception-action events. However, a real cognitive map is meant to include more than sequence-related knowledge; it actually implies the transformation of such knowledge into some more integrated representation. As TEC is not sufficiently developed to cover these kinds of transformation, it seems indeed promising to try connecting it with available ideas about how this might be achieved, as suggested by **Chown et al.**

### *R3.2. Stage models*

Some commentators (**Oriet et al.**; **Proctor & Vu**; **Sanders**) were skeptical about whether we did full justice to stage models of information processing and suggested that our discussion may have over-emphasized differences and under-emphasized commonalities between the stage approach and TEC. Interestingly, one of the major problems associated with stage accounts shows up in these commentaries themselves. According to both Donders' (1862) classical paper and Sternberg's (1969) foundation of the influential additive-factors logic (AFL), the term "stages" refers to cognitive processes. Thus, if two experimental factors produce a statistical interaction, one would be led to believe that they affect and, hence, are associated with, the same process. Logically speaking, identifying *processes* – the task AFL was developed for and, as **Sanders** points out, was very successful in – has no implications with respect to the attributes and the location of the *codes* these processes operate on. Assume, for instance, that additive-factors experiments would have revealed an interaction between factors presumably having to do with stimulus selection and factors presumably having to do with response selection. Given these findings, AFL would suggest that we think of stimulus and response selection being carried out by the same process, so that, if this process were capacity-limited, one may expect the concurrent selection of stimuli and responses to be difficult. So far so good. But would that have any implication as to what the selected codes look like and whether stimuli and responses are coded separately? Surely not, as the same process may work on codes anywhere in the system under analysis. And, indeed, it would be unfair to expect an approach developed to identify processes to deliver information about the architecture and functional location of the codes these processes operate on.

Among other things, it was the silence of stage approaches with respect to the latter that motivated us to develop TEC. For instance, we find it disappointing that about 70 years of research on the Psychological Refractory Period only allows us to confirm the assumption about some kind of bottleneck associated with the response-selection or "decision" stage that has existed since the very first studies onward (for an overview, see Pashler 1994) – without knowing *how* this stage really works and *in which way* it creates the bottleneck. True, localizing effects somewhere in the processing stream (or, better: stream of processes) provides an important starting point for analysis, but we are afraid that the actual analyses typically stopped right after the localizing was done. For these reasons, we are surprised to learn that TEC, with its emphasis on the architecture of, and interactions between codes, is not considered that different from existing stage models. However, close reading reveals that the assumed commonality results from a misinterpretation of AFL and the stage concept. For instance, **Sanders** writes that "the response selection stage has always been considered as a typical perceptual-motor interaction site," or, "according to stage theory 'late perceptual and early response products' typically reside in the response selection stage." It is obvious that in these sentences the term "stage" refers to functional locations or subsystems and, hence, used in a manner that is not covered by, and not consistent with AFL. Indeed, Sanders himself later admits that "stage theory has usually not had much to say about the contents and struc-



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

tures involved in stages." **Oriet et al.** also speak of perception and action as "systems" and thereby seem to refer to the location of codes rather than to the sequence of processes. We doubt that this is consistent with what the "classical model" they defend aims at.

We found ourselves similarly confused by **Proctor & Vu**'s view that TEC focuses "on that portion of information processing typically attributed to a central processing stage called response selection, response determination, response choice, or stimulus-response translation." For one thing, we were unable to find a single stage model where perceptual identification and action planning are handled by the same stage, whereas we had no problem in finding popular models that attribute these two processes to separate stages (for a selection, see Kornblum et al. 1990; Pashler 1994; Posner 1978; Proctor et al. 1990; Sanders 1983; Stoffels et al. 1989; Van Selst & Jolicoeur 1997; Welford 1952). Indeed, stage models would run into a lot of trouble if they really unified perceptual and action-planning *stages,* because this would predict all sorts of interactions between manipulations of "perceptual" and "action-related" factors that one does not (and presumably will not) find empirically; and in the *absence* of which the cited stage models are in fact founded! These kinds of problems do not arise from assuming common coding of perception and action plans, – which speaks for **Sanders**' suggestion to view stage theory and TEC as in some sense complementary.

### *R3.3. Ecological approaches*

The major goals of ecological approaches to human perception and action are to identify and characterize environmental (e.g., optical) variables that support behavior, and to describe the control laws or strategies employed to link variables and movements in an efficient, presumably task-dependent way. These goals stand in obvious contrast to those of modern cognitive psychology, which aims at understanding the relationship between cognitive processes and behavior. Both logically and empirically, there is of course some overlap between these approaches but it is true that ecological approaches emphasize the *What,* cognitive approaches the *How,* of information usage. We see no reason why either interest should not be scientifically legitimate and, in fact, see them as complementary in several ways. Yet, discussions between proponents of the two approaches are often dominated by the overarching idea that one approach must be more "correct" than the other (e.g., Meijer & Roth 1988; Weimer & Palermo 1974; 1982; Whiting 1984). The ecologically inspired commentaries on our target paper follow this tradition.

**Galantucci et al.** base their critical remarks on the observation that "in the tasks that support TEC, experimenters devise stimuli that can be described by sets of arbitrarily chosen, arbitrarily combined features (e.g., a letter is red or green; a rectangle is on the right or left side of a computer screen)," and they feel that "these sets are not up to the task of constituting percepts or action plans in nature." It appears that the validity of this description depends on how one conceives of "nature." TEC aims at accounting for how people master their daily life in and outside psychological labs in a world full of arbitrary combinations between features and between parameters of actions (see sect. R2.4). One may call this world "unnatural," and the perception-action codings it requires "unreal," but that does not save one from explaining how people can do rather well in perceiving and intentionally acting in it. In contrast to ecological models, TEC seems to be reasonably well prepared for such explaining. The remaining arguments of **Galantucci et al.** do not seem to contradict TEC: We neither assume that "percepts are . . . necessarily linear combinations of . . . features" (see sect. R2.4), nor find it unreasonable to assume that perception is in some sense "grounded" in action – an old idea explicitly endorsed in Hommel (1997; 1998a).

**Kim & Effken** criticize us for having "adopted the stimulus-response connection as . . . model for the perceiving-acting relation." If this were so, we would understand why Kim & Effken point out that "this model is fatally flawed" because "rarely is action interpretable as responses elicited by stimuli or in the presence of stimuli" and "neither is the perceiving-acting relation interpretable as a unidirectional effect of one upon the other." However, given that we took pains to reveal the *implausibility* of models based on stimulus-response connections (see sects. 2.1.1, 2.1.4, and 2.2.3.1 of our target article) and strongly argue *against* an uni-directional effect of stimuli on responses (a feature of TEC emphasized by **Proctor & Vu**), we find it very difficult to see where Kim & Effken are aiming at. Another issue they raise refers to the specification of future events. They feel that we "resort to cognitive representations of a common event code" because of our "incomplete understanding of ecological information." They then go on to point out that environmental information, such as time-to-contact (parameter tau), can specify the future state of affairs regarding both perception and action planning, so that resort to cognitive representations can be avoided. On the one hand, we agree with Kim & Effken that some aspects of sensorimotor coupling (the lower branch in Fig. 1) are likely to be controlled by the kind of information they have in mind. On the other hand, we have argued in section R2.2 that this kind of environmental control covers only a small part of what action control requires. Accordingly, as long as ecological approaches are unable to specify exactly *how* perceptual information, together with further disclosing "goals and effectivities," bring about actions such as preparing a cup of coffee, we find it not unreasonable to resort to such "mysterious or arcane explanatory devices" as codes in a human brain.

**Richardson & Michaels** have a somewhat more liberal attitude towards representations but they do not want them to be "mental" – a term that we did not find useful and, therefore, did not use. Instead, they ask for more consideration of the "specificational sense" of information and the codes it gives rise to, a theme also picked up by **Shaw & Wagman**. We agree with the general, in our view very healthy, ecological strategy of not complicating matters: if information is out there why not use it? But again, many aspects of our daily actions (even in long-jumping) cannot be specified by environmental information because this information is simply not "out there" (see sects. R2.2, R2.4). In contrast to ecological studies and models we do want to address the control of these aspects as well, which in the terminology of ecological psychology does require resort to "indicational information" as defined by Kelso and Kay (1987). Whether this makes TEC "void of explanatory power" should, in our view, be judged with respect to the empirical findings it correctly predicts, not by its fit with the aesthetical norms cultivated by a particular approach.


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

In contrast to the skeptical attitude one gathers from the ecologically motivated commentaries on TEC, we actually see no unsolvable problems in aligning our approach with recent developments in ecological theorizing. For instance, Warren (1998) suggests distinguishing between *model-based* and *information-based* control of action. The former is guided by representations that are informed by environmental states of affairs and frequently updated (i.e., by "event models" as defined by Stränger & Hommel 1996), but they can be used off-line to plan actions in the absence of suitable external information. The latter, in contrast, represents the type of on-line sensorimotor coupling that ecological accounts commonly focus on. Obviously, this distinction closely resembles Milner and Goodale's (1995) distinction of a perceptual and an action stream, and our own distinction between perception and action planning on the one hand, and sensorimotor processing on the other. Thus, we would think that sensorimotor processing streams can indeed be coupled to particular environmental variables (if present and sufficiently informative) by means of particular control laws or strategies (Fajen 2001); yet, it is the perception-action system addressed by TEC that is responsible, among other things, for selecting and implementing the most goal-satisfying strategy along the lines described elsewhere (Elsner & Hommel 2001; Hommel 1997; 1998a).

### *R3.4. Milner & Goodale's two-visual-pathways model*

**Pisella et al.** and **Westwood & Goodale** felt uneasy about our treatment, or the lack of it, of Milner and Goodale's (1995) distinction of a ventral processing stream underlying conscious visual perception and a dorsal stream driving "the transformation of visual information into spatially calibrated motor outputs." On the one hand, it is difficult to provide a fair comparison of a general framework for perception and action planning such as TEC, and an approach dealing with only one sensory modality and a very limited set of motor responses, such as manual pointing, aspects of grasping, and orienting one's hand. Obviously, the aims of the two approaches are different and so are both their level of specificity and the type of phenomena they refer to. Accordingly, TEC will be of little help in explaining many results that support Milner and Goodale's two-stream model, and the two-stream model will often face insurmountable problems if it comes to phenomena that we brought forward to support TEC. On the other hand, however, there are several aspects of Milner and Goodale's (1995) approach that, in our view, fit rather nicely with the picture drawn by TEC. According to their model, the dorsal pathway is strictly on-line and, hence, provides motor systems with up-to-date details about the environmental circumstances and the movement-related characteristics of goal-relevant objects. Such a pathway seems a perfect complement to what we think are the main attributes of perception and action planning: its (potential) off-line character, selectivity, and dependency on knowledge.

Accordingly, we tend to think of Milner and Goodale's dorsal stream as the sensorimotor pathway sketched in Figure 1. In contrast, the ventral stream of their model seems to share several attributes with the perception-action system TEC proposes. It is assumed to work off-line, to mediate associate learning and make use of its results, and to make sure that the animal's behavior meets emotional needs and social requirements. These features would make the ventral stream a perfect candidate for mediating perception and action planning along the lines of TEC. The major difference between our conception and that of Milner and Goodale (1995) and **Westwood & Goodale** lies in the question of what should be called "action planning" and "action control" – apart from the, in our view, secondary question: which neural structures are associated with conscious experience. Indeed, we doubt that it is appropriate to equate the "transformation of visual information into spatially calibrated motor outputs" with either term, a view that **Pisella et al.** seem to share. Somewhere in the system it needs to be determined what is transformed into what, when it is transformed, and to what end the transformations are carried out, all processes of action planning proper. If we understand Milner and Goodale (1995, especially p. 202) correctly, they would not suspect these decisions to be made within the ventral stream. This brings them to the unfortunate situation of having to admit that the real planning of an action – the processes that precede and control the planned sensorimotor transformations – actually take place outside the stream they devote to action but within a stream they devote to perception, a position **Westwood & Goodale** seem to endorse.

Things get even more problematic if we think of delayed actions, or actions in the absence of visual information. In such cases, Milner and Goodale (1995, p. 171) claim, a "stored percept of the object" from the ventral stream is used to feed the dorsal pathway, which allows for less, but often sufficiently precise performance. Although this is a reasonable assumption to make, in our view it further undermines the seemingly clear-cut distinction between a perceptual and an action pathway. A more transparent conceptualization that keeps many of the virtues of Milner and Goodale's approach seems to us the distinction between an on-line sensorimotor-transformation channel on the one hand and an off-line perception-action channel on the other, as proposed by **Pisella et al.** and TEC.

## R4. Concluding remarks

Thinking over our reply, we find ourselves in a somewhat ambiguous situation. At one level, we have defended the virtue of vagueness, but at the same time, at another level, we have offered a number of clarifications and specifications of what TEC is meant to be. This may be confusing at first glance, but actually we see no contradiction at all.

By emphasizing the importance of global, underspecified principles, we do not mean to say that the goal of science is underspecification – the less specified the better. In fact, we believe that science needs both weakly specified global principles and well-specified local theories in accordance with these principles. As regards global principles, we have two closing comments to offer (which, we believe, are of deep wisdom and thus cannot be contested . . . ): *First,* everything is a matter of degree. In order to fulfill their heuristic functions, global frameworks need to be underspecified *to an appropriate degree.* To be sure, being somewhat vague can be productive, but being too vague will certainly be detrimental. *Second,* everything is a matter of time. A global framework like TEC is not made for eternity. It is tailored to speak to the present state-of-the-art in the field, and we believe that it can play an important heuristic role *for a while.* When this while is over, TEC is bound to



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

die and fall into oblivion. However, we are confident that by then promising theoretical and experimental offspring will be emerging from it.

NOTE

**1.** We acknowledge the precedence of both Freud's *Instincts and Their Vicissitudes* (1915) and Neisser's *Stimulus Information and Its Vicissitudes* (a term Neisser borrowed from Freud for his monograph "Cognitive psychology," 1967).

## References

**Letters "a" and "r" appearing before authors' initials refer to target article and response, respectively.**


Abeles, M. (1991) *Corticonics: Neural circuits of the cerebral cortex.* Cambridge University Press. [aBH]

Ach, N. (1905) *Über die Willenstätigkeit und das Denken.* Vandenhoeck & Ruprecht. [aBH]

Adams, J. A. (1971) A closed-loop of motor learning. *Journal of Motor Behavior* 3:111–50. [aBH, AFS]

Albus, J. S. (1971) A theory of cerebellar function. *Mathematical Biosciences* 10:25–61. [PC]

Allport, D. A. (1987) Selection for action: Some behavioral and neurophysiological considerations of attention and action. In: *Perspectives on perception and action,* ed. H. Heuer & F. Sanders. Erlbaum. [aBH]

(1993) Attention and control: Have we been asking the wrong questions? A critical review of twenty-five years. In: *Attention and performance XIV,* ed. D. E. Meyer & S. Kornblum. MIT Press. [aBH]

Allport, D. A., Styles, E. A. & Hsieh, S. (1994) Shifting intentional set: Exploring the dynamic control of tasks. In: *Attention and performance XV,* ed. C. Umiltà & M. Moscovitch. MIT Press. [aBH, NM]

Allport, D. A., Tipper, S. P. & Chmiel, N. R. J. (1985) Perceptual integration and postcategorical filtering. In: *Attention and performance II,* ed. M. I. Posner & O. S. Marin. Erlbaum. [aBH]

Amazeen, E. & Turvey, M. T. (1996) Weight perception and the haptic size weight illusion are functions of the inertia tensor. *Journal of Experimental Psychology: Human Perception and Performance* 22:213–32. [MJR]

Andersen, R. A. (1988) The neurobiological basis of spatial cognition: Role of the parietal lobe. In: *Spatial cognition: Brain bases and development,* ed. J. Stiles-Davis, M. Kritchivsky & U. Belugi. Erlbaum. [aBH]

Anderson, J. R. (1993) *Rules of the mind.* Erlbaum. [aBH]

Arnoult, M. D. (1957) Stimulus predifferentiation: Some generalizations and hypotheses. *Psychological Bulletin* 54:339–50. [aBH]

Aschersleben, G. (2000a) Effects of training on the timing of simple repetitive movements. (submitted). [aBH]

(2000b) Knowledge of results and the timing of actions. (Paper No. 1/2000). Max-Planck-Institut für Psychologische Forschung. [aBH]

(in press) Temporal control of movements in sensorimotor synchronization. *Brain and Cognition.* [aBH]

Aschersleben, G., Gehrke, J. & Prinz, W. (2001) Tapping with peripheral nerve block: A role for tactile feedback in the timing of movements. *Experimental Brain Research* 136:331–39. [aBH]

(in press) A psychophysical approach to action timing. In: *Psychophysics beyond sensation: Laws and invariants in human cognition,* ed. C. Koernbach, E. Schröger & H. Müller. Erlbaum. [aBH]

Aschersleben, G. & Müsseler, J. (1999) Dissociations in the timing of stationary and moving stimuli. *Journal of Experimental Psychology: Human Perception and Performance* 25:1709–20. [aBH]

Aschersleben, G. & Prinz, W. (1995) Synchronizing actions with events: The role of sensory information. *Perception and Psychophysics* 57:305–17. [arBH]

(1997) Delayed auditory feedback in synchronization. *Journal of Motor Behavior* 29:35–46. [aBH]

Aschersleben, G., Stenneken, P., Cole, J. & Prinz, W. (2002) Timing mechanisms in sensorimotor synchronization. In: *Attention and performance XIX: Common mechanisms in perception and action,* ed. W. Prinz & B. Hommel. Oxford University Press. [aBH]

Atkeson, C. G. (1989) Learning arm kinematic and dynamics. *Annual Review of Neuroscience* 12:157–83. [DAR]

Atkinson, R. C. & Shiffrin, R. M. (1968) Human memory: A proposed system and its control processes. In: *The psychology of learning and motivation,* ed. K. W. Spence & J. D. Spence. Academic Press. [MOB]

Babcock, M. K. & Freyd, J. J. (1988) The perception of dynamic information in static handwritten forms. *American Journal of Psychology* 101:11–130. [aBH]

Ballard, D. H., Hayhoe, M. M., Pook, P. K. & Rao, R. P. N. (1997) Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences* 20(4):723–67. [DCR]

Bandura, A. (1977) *Social learning theory.* Prentice Hall. [aBH]

Bandura, A. & Walters, R. H. (1963) *Social learning and personality development.* Holt, Rinehart, & Winston. [aBH]

Bargh, J. A. (1996) Automaticity in social psychology. In: *Social psychology: Handbook of basic principles,* ed. E. T. Higgins & A. W. Kruglanski. Guilford Press. [aBH]

Bargh, J. A. & Barndollar, K. (1996) Automaticity in action: The unconscious as repository of chronic goals and motives. In: *The psychology of action: Linking cognition and motivation to behavior,* ed. P. M. Gollwitzer & J. A. Bargh. Guilford Press. [aBH]

Bargh, J. A. & Gollwitzer, P. M. (1994) Environmental control of goal-directed action: Automatic and strategic contingencies between situations and behavior. *Nebraska Symposium on Motivation* 41:71–124. [rBH]

Barker, R. G. (1963) The stream of behavior as an empirical problem. In: *The stream of behavior,* ed. R. G. Barker. Appleton. [aBH]

Barker, R. G. & Wright, H. F. (1954) *Midwest and its children: The psychological decay of an American town.* Row, Peterson. [JMZ]

Barsalou, L. W. (1999a) Perceptual symbol systems. *Behavioral and Brain Sciences* 22(4):577–660. [DCR]

(1999b) Language comprehension: Archival memory or preparation for situated action? *Discourse Processes* 28(1):61–80. [DCR]

Basso, D. & Bisiacchi, P. S. (1999) Il dilemma del commesso viaggiatore: Uno studio computerizzato. *A. I. P. Associazione Italiana di Psicologia, Congresso Nazionale della Sezione Psicologia Sperimentale, Capri Novantanove,* 155–57. A. I. P. (Associazione Italiana di Psicologia Press). [MOB]

Bechara, A., Damasio, H., Tranel, D. & Anderson, S. W. (1998) Dissociation of working memory from decision making within the human prefrontal cortex. *Journal of Neuroscience* 18:428–37. [PC]

Beck, J. (1982) Textural segmentation. In: *Organization and representation in perception,* ed. J. Beck. Erlbaum. [aBH]

Beckers, T. & De Houwer, J. (2000) Persistent integration of action-effect codes in the cognitive representation of actions: A conditioning perspective. Poster presented at Attention & Performance XIX, Kloster Isee, Germany. [rBH]

Bender, S., Müller, B., Oades, R. D. & Sartory, G. (2001) Conditioned blocking and schizophrenia: A replication and study of the role of symptoms, age, onset-age of psychosis and illness-duration. *Schizophrenia Research* 49:157–70. [RDO]

Bernstein, N. A. (1967) *The coordination and regulation of movements.* Pergamon Press. [N-GK]

Bock, J. K. (1986) Syntactic persistence in language production. *Cognitive Psychology* 18:355–87. [RJH]

(1996) Language production: Methods and methodologies. *Psychonomic Bulletin and Review* 3:395–421. [RJH]

Branigan, H. P., Pickering, M. J. & Clelland, A. A. (2000) Syntactic co-ordination in dialogue. *Cognition* 75:B1-B13. [RJH]

Brewer, B. (1993) The interpretation of spatial vision and action. In: *Spatial representation. Problems in philosophy and psychology,* ed. N. Eilan, R. McCarthy & B. Brewer. Blackwell. [aBH]

Bridgeman, B. (1983) Mechanisms of space constancy. In: *Spatially oriented behavior,* ed. A. Hein & M. Jeannerod. Springer. [aBH]

Bridgeman, B., Kirch, M. & Sperling, A. (1981) Segregation of cognitive and motor aspects of visual function using induced motion. *Perception and Psychophysics* 29:336–42. [aBH]

Broadbent, D. E. (1958) *Perception and communication.* Pergamon Press. [aBH, JI]

(1971) *Decision and stress.* Academic Press. [aBH]

(1993) A word before leaving. In: *Attention and performance XIV,* ed. D. E. Meyer & S. Kornblum. Erlbaum. [rBH]

Brooks, R. A. (1986) A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation* RA-2:14–23. [JB]

Bruno, N. (2001) When does action resist visual illusions? *Trends in Cognitive Sciences* 5:379–82. [rBH]

Brunswik, E. (1944) Distal focussing of perception. Size constancy in a representative sample of situations. *Psychological Monographs* 56(1). [aBH]

Bryson, J. & Stein, L. A. (2001) Modularity and design in reactive intelligence. In: *Proceedings of the 17th International Joint Conference on Artificial Intelligence (IJCAI01), Seattle.* Morgan Kaufmann. [JB]

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G. & Freund, H.-J. (2001) Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience* 13(2):400–404. [TC]

Bundesen, C. (1990) A theory of visual attention. *Psychological Review* 97:523–47. [rBH]




https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

Burnod, Y, Baraduc, P., Battaglia-Mayer, A., Guignon, E., Koechlin, E., Ferraina, S., Lacquaniti, F. & Caminiti, R. (1999) Parieto-frontal coding of reaching: An integrated framework. *Experimental Brain Research* 129:325–46. [SV]

Cadwallader, M. (1975) A behavioral model of consumer spatial decision making. *Economic Geography* 51:339–49. [MOB]

Caessens, B. & Vandierendonck, A. (2001) Inhibition and blindness to response compatible stimuli: A reappraisal. (submitted). [WK]

Caminiti, R., Ferraina, S. & Battaglia-Mayer, A. (1998) Visuomotor transformations: Early cortical mechanisms of reaching. *Current Opinion in Neurobiology* 8:753–61. [SV]

Campbell, B. A., Krauter, E. E. & Wallace, J. E. (1980) Animal models of aging: Sensory-motor and cognitive function in the aged rat. In: *The psychology of aging: Problems and perspectives,* ed. W. Stein. Elsevier/North Holland. [HRD]

Cantor, J. H. (1965) Transfer of stimulus pretraining to motor paired-associate and discrimination learning tasks. In: *Advances in child development and behavior, vol. 2,* ed. L. Lipsitt & C. C. Spiker. Academic Press. [aBH]

Carey, D. P. (2001) Do action systems resist visual illusion? *Trends in Cognitive Sciences* 5(3):109–113. [rBH]

Carlson, N. R. (2000) *Physiology of behavior.* Allyn and Bacon. [aBH]

Carpenter, R. H. S. (1988) *Movement of the eyes.* Pion. [aBH]

Carpenter, W. B. (1874) *Principles of mental physiology, with their applications to the training and discipline of the mind and the study of its morbid conditions.* Appleton. [aBH]

Chaminade, T., Meary, D., Orliaguet, J.-P. & Decety, J. (in press) Is perceptual anticipation a motor simulation? A PET study. *NeuroReport.* [TC]

Chaminade, T., Meltzoff, A. N. & Decety, J. (in press) Does the end justifies the means? A PET exploration of the mechanisms involved in imitation. *Neuroimage.* [TC]

Chase, W. G. & Simon, H. A. (1973) The mind's eye in chess. In: *Visual information processing,* ed. W. G. Chase. Academic Press. [PCRL]

Cheng, P. C.-H. (1996) Scientific discovery with law encoding diagrams. *Creativity Research Journal* 9:145–62. [PCRL]

(1998) A framework for scientific reasoning with law encoding diagrams: Analysing protocols to assess its utility. In: *Proceedings of the Twentieth Annual Meeting of the Cognitive Science Society,* pp. 232–35. Erlbaum. [PCRL]

Cheng, P. C.-H., McFadzean, J. & Copeland, L. (2001) Drawing on the temporal signature of induced perceptual chunks. In: *Proceedings of the Twenty-third Annual Meeting of the Cognitive Science Society, Edinburgh, UK,* pp. 200–205, ed. J. D. Moore & K. Stenning. Erlbaum. [PCRL]

Cherry, E. C. (1953) Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America* 25:975–79. [aBH]

Chown, E., Kaplan, S. & Kortenkamp, D. (1995) Prototypes, location, and associative networks (PLAN): Towards a unified theory of cognitive mapping. *Cognitive Science* 19:1–51. [EC]

Chur, L., Spengler, F., Jürgens, M. & Dinse, H. R. (1996) Environmental enrichment counteracts decline of sensorimotor performance and deterioration of cortical organization in aged rats. *Society for Neuroscience Abstracts* 22:102. [HRD]

Cisek, P. (1999) Beyond the computer metaphor: Behaviour as interaction. *Journal of Consciousness Studies* 6:125–42. [PC]

(in press) Embodiment is all in the head. *Behavioral and Brain Sciences.* [PC]

Clark, A. (1997) *Being there: Putting brain, body, and the world together again.* MIT Press. [DCR]

Cohen, C. E. & Ebbesen, E. B. (1979) Observational goals and schema activation: A theoretical framework for behavior perception. *Journal of Experimental Social Psychology* 15:305–29. [aBH]

Cohen, J. D., Braver, T. S. & O'Reilly, R. C. (1998) A computational approach to prefrontal cortex, cognitive control, and schizophrenia: Recent developments and current challenges. In: *The prefrontal cortex: Executive and cognitive functions,* ed. A. C. Roberts, T. W. Robbins & L. Weiskrantz. Oxford University Press. [rBH]

Cohen, J. D., Dunbar, K. & McClelland, J. L. (1990) On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review* 97:332–61. [rBH]

Colby, C. L. (1998) Action-oriented spatial reference frames in the cortex. *Neuron* 20:15–24. [aBH]

Conditt, M. A., Gandolfo, F. & Mussa-Ivaldi, F. A. (1997) The motor system does not learn the dynamics of the arm by rote memorization of past experience. *Journal of Neurophysiology* 78:554–60. [DAR]

Constantinidis, C., Franowicz, M. N. & Goldman-Rakic, P. S. (2001) The sensory nature of mnemonic representation in the primate prefrontal cortex. *Nature Neuroscience* 4:311–16. [PC]

Conway, B. A., Halliday, D. M., Farmer, S. F., Shahani, U., Maas, P., Weir, A. I. & Rosenberg, J. R. (1995) Synchronization between motor cortex and spinal motoneural pool during the performance of a maintained motor task in man. *Journal of Physiology* 489:917–24. [aBH]

Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., Linenweber, M. R., Petersen, S. E., Raichle, M. E., Van-Essen, D. C. & Shulman, G. L. (1998) A common network of functional areas for attention and eye movements. *Neuron* 21(4):761–73. [JI]

Cowan, N. (2001) The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences* 24:87–186. [PCRL]

Cowey, A. (1985) Aspects of cortical organization related to selective attention and selective impairments of visual perception. In: *Attention and performance XI,* ed. M. I. Posner & O. S. M. Marin. Erlbaum. [aBH]

Craighero, L., Fadiga, L., Rizzolatti, G. & Umiltà, C. (1999) Movement for perception: A motor-visual attentional effect. *Journal of Experimental Psychology: Human Perception and Performance* 25(6):498–511. [rBH, WK, SV]

Craik, K. J. W. (1943) *The nature of explanation.* Cambridge University Press. [EC]

Cutting, J. E. & Proffitt, D. R. (1981) Gait perception as an example of how we may perceive events. In: *Intersensory perception and sensory integration,* ed. R. Walk & H. Pick. Plenum Press. [aBH]

Damasio, A. R. (1994) *Descartes' error.* Putnam. [rBH]

Davies, P. (1988) *The cosmic blueprint.* Simon & Schuster. [RES]

Decety, J. (in press) Is there such a thing as a functional equivalence between imagined, observed, and executed action? In: *The imitative mind: Development, evolution, and brain bases,* ed. A. N. Meltzoff & W. Prinz. Cambridge University Press. [aBH]

Decety, J., Chaminade, T., Grèzes, J. & Meltzoff, A. N. (in press) A PET exploration of the neural mechanisms involved in reciprocal imitation. *Neuroimage.* [TC]

Decety, J. & Grèzes, J. (1999) Neural mechanisms subserving the perception of human actions. *Trends in Cognitive Neuroscience* 3:172–78. [aBH]

Decety, J., Grèzes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F. & Fazio, F. (1997) Brain activity during observation of actions: Influence of action content and subject's strategy. *Brain* 120(10):1763–77. [TC]

De Groot, A. D. & Gobet, F. (1996) *Perception and memory in chess: Heuristics of the professional eye.* Van Gorcum. [PCRL]

De Jong, R. (1993) Multiple bottlenecks in overlapping task performance. *Journal of Experimental Psychology: Human Perception and Performance* 19:965–80. [aBH]

(2000) An intention-activation account of residual switch costs. In: *Control of cognitive processes: Attention and performance XVIII,* ed. S. Monsell & J. Driver. MIT Press. [rBH]

De Jong, R., Liang, C-C. & Lauber, E. (1994) Conditional and unconditional automaticity: A dual-process model of effects of spatial, stimulus-response correspondence. *Journal of Experimental Psychology: Human Perception and Performance* 20:731–50. [aBH]

De Jong, R. & Sweet, J. B. (1994) Preparatory strategies in overlapping-task performance. *Perception and Psychophysics* 55:142–51. [aBH]

Della Malva, C. L., Stuss, D. T., D'Alton, J. & Willmer, J. (1993) Capture errors and sequencing after frontal brain lesion. *Neuropsychologia* 31(4):363–72. [MOB]

Derwort, A. (1938) Untersuchungen über Zeitverlauf figurierter Bewegungen bei Menschen. *Pflügers Archiv für die gesammte Physiologie* 240:261–75. [AFS]

Descartes, R. (1664) *L'homme.* Theodore Girard. [aBH]

Desimone, R. & Ungerleider, L. G. (1989) Neural mechanisms of visual processing in monkeys. In: *Handbook of neuopsychology, vol. 2,* ed. F. Boller & J. Grafman. Elsevier. [aBH]

Desmurget, M. & Grafton, S. T. (2000) Forward modelling allows feedback control for fast reaching movements. *Trends in Cognitive Sciences* 4(11):423–31. [LP]

Desmurget, M., Grea, H., Grethe, J. S., Prablanc, C., Alexander, G. E. & Grafton, S. T. (2001) Functional anatomy of nonvisual feedback loops during reaching: A positron emission tomography study. *Journal of Neuroscience* 21(8):2919–28. [LP]

Deubel, H. & Schneider, W. X. (1996) Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research* 36(12):1827–37. [JI]

Deubel, H., Schneider, W. X. & Paprotta, I. (1998) Selective dorsal and ventral processing: Evidence for a common attentional mechanism in reaching and perception. *Visual Cognition* 5(1–2):81–107. [JI]

Deutsch, J. A. & Deutsch, D. (1963) Attention: Some theoretical considerations. *Psychological Review* 20: 80–90. [aBH]

Dewey, J. (1896) The reflex arc concept in psychology. *Psychological Review* 3:357–70. [aBH]

DeYoe, E. A. & Van Essen, D. C. (1988) Concurrent processing streams in monkey visual cortex. *Trends in Neuroscience* 11:219–26. [aBH]

Diedrichsen, J., Hazeltine, E., Kennerley, S. & Ivry, R. B. (in press) Moving to directly cued locations abolishes spatial interference during bimanual actions. *Psychological Science.* [JD]



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

Diedrichsen, J., Ivry, R. B., Cohen, A. & Danziger, S. (2000) Asymmetries in a unilateral flanker task depend on the direction of the response: The role of attentional shift and perceptual grouping. *Journal of Experimental Psychology: Human Perception and Performance* 26(1):113–26. [JD, rBH]

Dinse, H. R. (1994) A time-based approach towards cortical functions: Neural mechanisms underlying dynamic aspects of information processing before and after postontogenetic plastic processes. *Physica D* 75:129–50. [HRD]

(2001) The aging brain. In: *The neurosciences at the turn of the century. Proceedings of the Fourth Meeting of the German Neuroscience Society, vol. 1,* ed. N. Elsner & G. W. Kreutzberg. Thieme. [HRD]

Dinse, H. R. & Krüger, K. (1994) Timing of processing along the visual pathway in the cat. *NeuroReport* 5:893–97. [HRD]

Dinse, H. R. & Merzenich, M. M. (2001) Adaptation of inputs in the somatosensory system. In: *Perceptual learning,* ed. M. Fahle & T. Poggio. MIT Press. (in press). [HRD]

Dinse, H. R., Zepka, R. F., Jürgens, M., Godde, B., Hilger, H. & Berkefeld, T. (1995) Age-dependent changes of cortical and thalamic representations revealed by optical imaging and electrophysiological mapping techniques: Evidence for degenerative and use-disuse-dependent processes. *Homeostasis in Health and Disease* 36:S1, 49. [HRD]

Di Pellegrino, G., Fadiga, L., Fogassi, V., Gallese, V. & Rizzolatti, G. (1992a) Understanding motor events: A neurophysiological study. *Experimental Brain Research* 91:176–80. [aBH, CO]

Di Pellegrino, J. W., Klatzky, R. L. & McCloskey, B. P. (1992b) Time course of preshaping for functional responses to objects. *Journal of Motor Behaviour* 21:307–16. [aBH]

Domjan, M., Cusato, B. & Villareal, R. (2000) Pavlovian feed-forward in the control of social behavior. *Behavioral and Brain Sciences* 23:235–82. [MOB]

Donders, F. C. (1862) Über die Schnelligkeit psychischer Processe. *Archiv für Anatomie und Physiologie,* 657–81. [rBH]

(1868a/1969) Die Schnelligkeit psychischer Processe. *Reichert's und Dubois-Reymond's Archiv für Anatomie, Physiologie und Wissenschaftliche Medizin,* 657–81. (1969 English edition : On the speed of mental processes. *Acta Psychologica* 30:412–31.) [aBH]

(1868b) Over de snelheid van psychische processen. Onder-zoekingen gedaan in het physiologisch laboratorium van de Utrechtse Hoogeschool. *Tweede Reeks* 11:92–130. [AFS]

Dretske, F. I. (1981) *Knowledge and the flow of information.* MIT Press. [aBH, RGM]

(1986) Misrepresentation. In: *Belief: Form, content, and function,* ed. R. Bogdan. Oxford University Press. [RGM]

(1988) *Explaining behavior.* MIT Press. [RGM]

(1995) *Naturalizing the mind.* MIT Press. [RGM]

Duhamel, J.-R., Colby, C. L. & Goldberg, M. E. (1992) The updating of the representation of visual space in parietal cortex by intended eye movements. *Science* 255:90–92. [PC]

Duncan, J. (1993) Selection and input and goal in the control of behavior. In: *Selection, awareness, and control,* ed. A. Baddeley & L. Weiskrantz. Clarendon Press. [rBH]

Duncan, J. & Humphreys, G. W. (1989) Visual search and stimulus similarity. *Psychological Review* 96:433–58. [rBH, WK, LP]

(1992) Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance* 18:578–88. [LP]

Eckhorn, R., Frien, A., Bauer, R., Woelbern, T. & Kehr, H. (1993) High frequency (60–90Hz) oscillations in primary visual cortex of awake monkey. *NeuroReport* 4:243–46. [aBH]

Edelman, G. M. (1996) *Bright air, brilliant fire.* Basic Books. [RES]

Egeth, H. E. (1966) Parallel vs. serial processes in multidimensional stimulus discrimination. *Perception and Psychophysics* 1:245–52. [aBH]

Ehrenstein, W. H. (1994) The Simon effect and visual motion. *Psychological Research* 56:163–69. [aBH]

Ellis, R. & Tucker, M. (2000) Micro-affordances: The potentiation of components of action by seen objects. *British Journal of Psychology* 91(4):451–71. [TC]

Elsner, B. (2000) *Der Erwerb kognitiver Handlungsrepräsentationen.* Wissenschaftlicher Verlag. [aBH]

Elsner, B. & Hommel, B. (2000) Effect anticipation and action control. *Journal of Experimental Psychology: Human Perception and Performance* 27:229–40. [arBH]

Engel, A. K., König, P., Kreiter, A. K. & Singer, W. (1991) Interhemispheric synchronization of oscillatory neuronal responses in cat visual cortex. *Science* 252:1177–79. [aBH]

Enright, S. J. & Beech, A. R. (1993) Further evidence of reduced cognitive inhibition in obsessive-compulsive disorder. *Personality and Individual Differences* 14:387–95. [RDO]

Epstein, W. (1973) The process of "taking-into-account" in visual perception. *Perception* 2:267–85. [aBH]

Eskander, E. N. & Assad, J. A. (1999) Dissociation of visual, motor and predictive signals in parietal cortex during visual guidance. *Nature Neuroscience* 2:88–93. [PC]

Fabre-Thorpe, M., Delorme, A., Marlot, C. & Thorpe, S. (2001) A limit to the speed of processing in ultra-rapid categorization of novel natural scenes. Journal of *Cognitive Neuroscience* 13(2):171–80. [MOB]

Fadiga, L., Fogassi, L., Gallese, G. & Rizzolatti, G. (2000) Visuomotor neurons: Ambiguity of the discharge or "motor" perception? *International Journal of Psychophysiology* 35:165–77. [PC, aBH]

Fadiga, L., Fogassi, L., Pavesi, G. & Rizzolatti, G. (1995) Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology* 73:2608–11. [aBH]

Fagot, C. (1994) *Chronometric investigations of task switching.* Ph. D. thesis, University of California, San Diego. [NM]

Fahen, B. R. (2001) Steering towards a goal by equalizing taus. *Journal of Experimental Psychology: Human Perception and Performance* 27:229–40. [rBH]

Fehrer, E. & Raab, E. (1962) Reaction time to stimuli masked by metacontrast. *Journal of Experimental Psychology* 63:143–47. [aBH]

Feigenbaum, E. A. & Simon, H. A. (1984) EPAM-like models of recognition and learning. *Cognitive Science* 8:305–36. [PCRL]

Felleman, D. J. & Van Essen, D. C. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex* 1:1–47. [HRD]

Ferlazzo, F., Padovani, T., Couyoumdijan, A. & Olivetti Belardinelli, M. (2001) Head-centered meridian effect on auditory spatial attention orienting. *The Quarterly Journal of Experimental Psychology: A.* (in press). [MOB]

Ferraro, F. R., Jones, D. & McManus, G. (1996) Relationship of Maudsley obsessive-compulsive inventory and negative priming: Some preliminary evidence. *Brain and Cognition* 32:284–85. [RDO]

Fischer, M. H. (1999) An investigation of attention allocation during sequential eye movement tasks. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology* 52A(3):649–77. [JI]

Fitts, P. M. & Seeger, C. M. (1953) S-R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology* 46:199–210. [aBH, RWP]

Folk, C. L., Remington, R. W. & Johnston, J. C. (1992) Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance* 18(4):1030–44. [rBH, JI]

Fowler, C. A. & Turvey, M. T. (1982) Observational perspective and descriptive level in perceiving and acting. In: *Cognition and the symbolic processes, vol. II,* ed. W. B. Weimer & D. S. Palermo. Erlbaum. [aBH]

Fox, E. (1995) Negative priming form ignored distractors in visual selection: A review. *Psychonomic Bulletin and Review* 2:145–73. [aBH]

Fraisse, P. (1980) Les synchronisations sensori-motrices aux rythmes [The sensorimotor synchronization of rhythms]. In: *Anticipation et comportement,* ed. J. Requin. Centre National. [aBH]

Freeman, A. & Oster, G. L. (1998) Cognitive therapy and depression. In: *International handbook of cognitive and behavioral treatments for psychological disorders,* ed. V. E. Caballo. Pergamon/Elsevier. [MOB]

Freud, S. (1915/1956) Instincts and their vissicitudes. In: *Collected papers of Sigmund Freud. Vol. IV.* Hogarth Press. (1956 edition). [rBH]

Freyd, J. J. (1987) Dynamic mental representations. *Psychological Review* 94:427–38. [aBH]

Frith, C. D. (1992) *The cognitive neuropsychology of schizophrenia.* Erlbaum. [RDO]

Frith, C. D., Blakemore, S.-J. & Wolpert, D. (2000) Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society of London B* 355:1771–88. [LP]

Fuster, J. M. (2000) Prefrontal neurons in networks of executive memory. *Brain Research Bulletin* 52:331–36. [PC]

Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. (1996) Action recognition in the premotor cortex. *Brain* 119:593–609. [aBH, RDO]

(2002) Action representation and the inferior parietal lobule. In: *Attention and performance, vol. XIX: Common mechanisms in perception and action,* ed. W. Prinz & B. Hommel. Oxford University Press. [rBH]

Gallistel, C. R. (1990) *The organization of learning.* MIT Press. [JH]

Garrett, M. F. (2000) Remarks on the architecture of language processing systems. In: *Language and brain,* ed. Y. Grodzinsky, L. Shapiro & D. Swinney. Academic Press. [RJH]

Gehrke, J. (1995) *Sensorimotor synchronization: The intensity of afferent feedback affects the timing of movements* (Paper 15/1995). Max Planck Institute for Psychological Research. [aBH]

(1996) Afferente Informationsverarbeitung und die Synchronisation von Ereignissen [Processing of afferent information and the synchronization of events]. Unpublished dissertation, Munich, Germany. [aBH]

Gehrke, J. & Hommel, B. (1998) The impact of exogenous factors on spatial coding in perception and memory. In: *Spatial cognition: An interdisciplinary*



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

*approach to representing and processing spatial knowledge,* ed. C. Freksa, C. Habel & K. F. Wender. Springer. [rBH]
Gewirtz, J. L. & Stingle, K. G. (1968) Learning of generalized imitation as the basis for identification. *Psychological Review* 75:374–97. [aBH]
Gibson, E. J. (1969) *Principles of perceptual learning and development.* Prentice-Hall. [aBH, AFS]
(1977) The theory of affordances. In: *Perceiving, acting, and knowing,* ed. R. E. Shaw & J. Bransford. Erlbaum. [aBH, AFS]
Gibson, E. J. & Levin, H. (1975) *The psychology of reading.* MIT Press. [AFS]
Gibson, J. J. (1950) *The perception of the visual world.* Houghton Mifflin. [aBH, N-GK]
(1966) *The senses considered as perceptual systems.* Houghton Mifflin. [arBH, N-GK, MJR, RES]
(1979) *The ecological approach to visual perception.* Houghton Mifflin. [JH, arBH, N-GK, MJR, RES]
Gibson, J. J. & Gibson, E. J. (1955) Perceptual learning: Differentiation or enrichment? *Psychological Reviews* 62:32–41. [aBH]
Glenberg, A. (1997) What memory is for. *Behavioral and Brain Sciences* 20(1):1–55. [DCR]
Gobet, F., Lane, P. C. R., Croker, S. J., Cheng, P. C.-H., Jones, G., Oliver, I. & Pine, J. M. (2001) Chunking mechanisms in human learning. *Trends in Cognitive Science* 5:236–43. [PCRL]
Gobet, F. & Jansen, P. (1994) Towards a chess program based on a model of human memory. In: *Advances in computer chess,* ed. H. J. van den Herik, I. S. Herschberg & J. W. Uiterwijk. University of Linburgh Press. [PCRL]
Gobet, F. & Simon, H. A. (2000) Five seconds or sixty? Presentation time in expert memory. *Cognitive Science* 24:651–82. [PCRL]
Godde, B., Spengler, G. & Dinse, H. R. (1996) Associative pairing of tactile stimulation induces somatosensory cortical reorganization in rats and humans. *NeuroReport* 8:281–85. [HRD]
Godde, B., Stauffenberg, B., Spengler, F. & Dinse, H. R. (2000) Tactile coactivation induced changes in spatial discrimination performance. *Journal of Neuroscience* 20:1597–604. [HRD]
Gollwitzer, P. M. (1996) The volitional benefits of planning. In: *The psychology of action: Linking cognition and motivation to behavior,* ed. P. M. Gollwitzer & J. A. Bargh. Guilford Press. [aBH]
Goodale, M. A. (1997) Visual routes to perception and action in the cerebral cortex. In: *Handbook of neuropsychology, vol. 11, section 16: Action and cognition,* ed. M. Jeannerod. Series editors: F. Boller & J. Grafman. Elsevier. [SV]
(1998) Visuomotor control: Where does vision end and action begin? *Current Biology* 8:489–91. [DAW]
Goodale, M., Pélisson, D. & Prablanc, C. (1986) Large adjustments in visually guided reaching do not depend on vision of the hand or perception of target displacement. *Nature* 320:748–50. [rBH]
Goodwin, B. (1994) *How the leopard changed its spots: The evolution of complexity.* Charles Scribner's. [RES]
Gopher, D. & Sanders, A. F. (1984) S-Oh-R: Oh stages! Oh resources! In: *Cognition and motor behavior,* ed. W. Prinz & A. F. Sanders. Springer. [AFS]
Gordon, P. C. (1990) Perceptual motor processing in speech. Stimulus-response compatibility: An integrated perspective. In: *Stimulus-response compatibility: An integrated perspective,* ed. R. W. Proctor & T. G. Reeve. North-Holland. [aBH]
Gordon, R. D. & Irwin, D. E. (1996) What's in an object file? Evidence from priming studies. *Perception and Psychophysics* 58:1260–77. [aBH]
(1998) Eye movements, attention and trans-saccadic memory. *Visual Cognition* 5:127–55. [aBH]
(2000) The role of physical and conceptual properties in preserving object continuity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26:136–50. [aBH]
Gould, S. J. & Lewontin, R. C. (1979) The spandrels of San Marco and the Panglossian paradigm: A critique of the adaptationist programme. *Proceedings of the Royal Society of London B* 205:581–98. [RES]
Grafton, S. T., Arbib, M. A., Fadiga, L. & Rizzolatti, G. (1996) Localization of grasp representations in humans by positrone emission tomography. 2. Observation compared with imagination. *Experimental Brain Research* 112:103–11. [aBH]
Gray, C. M., Engel, A. K., Koenig, P. & Singer, W. (1992) Synchronization of oscillatory neuronal responses in cat striate cortex: Temporal properties. *Visual Neuroscience* 8(4):337–47. [aBH]
Graziano, M. S. A. & Gandhi, S. (2000) Location of the polysensory zone in the precentral gyrus of anesthetized monkeys. *Experimental Brain Research* 135(2):259–66. [JB]
Graziano, M. S. A., Taylor, C. S. R. & Moore, T. (2001) Brain mechanisms for monitoring and controlling the body. (personal communication). [JB]
Greene, P. H. (1982) Why is it easy to control your arms? *Journal of Motor Behavior* 14:260–86. [aBH]
(1988) The organization of natural movement. *Journal of Motor Behavior* 20:180–85. [aBH]
Greenwald, A. G. (1970) Sensory feedback mechanisms in performance control: With special reference to the ideomotor mechanism. *Psychological Review* 77:73–99. [aBH, AFS]
(1972) On doing two things at once: Time sharing as a function of ideomotor compatibility. *Journal of Experimental Psychology* 94:52–57. [aBH]
Greenwald, A. G. & Shulman, H. G. (1973) On doing two things at once: II. Elimination of the psychological refractory period effect. *Journal of Experimental Psychology* 101(1):70–76. [WK]
Grèzes, J. & Decety, J. (2001) Functional anatomy of execution, mental simulation, observation and verb generation of actions: A meta-analysis. *Human Brain Mapping* 12(1):1–19. [TC]
(2002) Does visual perception of objects afford action. Evidence from a neuroimaging study. *Neuropsychologia* 40:212–22. [TC]
Grossberg, S. (1969) On learning of spatiotemporal patterns by networks with ordered sensory and motor components. I. Excitatory components of the cerebellum. *Studies in Applied Mathematics* 48:105–32. [PC]
Grossberg, S. & Kuperstein, M. (1989) *Neural dynamics of adaptive sensory-motor control: Expanded edition.* Pergamon Press. [aBH]
Guiard, Y. (1983) The lateral coding of rotations: A study of the Simon effect with wheel-rotation responses. *Journal of Motor Behavior* 15:331–42. [aBH]
Gurney, K., Prescott, T. J. & Redgrave, P. (1998) The basal ganglia viewed as an action selection device. In: *The Proceedings of the International Conference on Artificial Neural Networks, Skovde, Sweden.* [JB]
Haber, R. N. (1969) Introduction. In: *Information-processing approaches to visual perception,* ed. R. N. Haber. Holt, Rinehart, & Winston. [aBH]
Haig, A. R., Gordon, E., De Pascalis, V., Meares, R. A., Bahramali, H. & Harris, A. (2000) Gamma activity in schizophrenia: Evidence of impaired network binding? *Clinical Neurophysiology* 111:1461–68. [RDO]
Hall, G. (1991) *Perceptual learning and associative learning.* Clarendon Press. [aBH]
Harless, E. (1861) Der Apparat des Willens. *Zeitschrift für Philosophie und philosophische Kritik* 38:50–73. [rBH]
Hartston, H. J. & Swerdlow, N. R. (1999) Visiospatial priming and Stroop performance in patients with obsessive-compulsive disorder. *Neuropsychology* 13:447–57. [RDO]
Hartsuiker, R. J. & Kolk, H. H. J. (2001) Error monitoring in speech production: A computational test of the perceptual loop theory. *Cognitive Psychology* 42:113–57. [RJH]
Hayes-Roth, B. & Hayes-Roth, F. (1979) A cognitive model of planning. *Cognitive Science* 3:275–310. [MOB]
Hazeltine, E., Diedrichsen, J., Kennerley, S. & Ivry, R. B. (in preparation) Bimanual interference during reaching movements is primarily related to response selection, not the specification of motor parameters. [JD]
Hebb, D. O. (1949) *The organization of behavior.* Wiley. [EC]
Hecht, H., Vogt, S. & Prinz, W. (2001) Motor learning enhances perceptual judgement: A case for action-perception transfer. *Psychological Research/Psychologische* Forschung 65:3–14. [SV]
Heider, F. (1926/1959) Thing and medium. *Psychological Issues* 1959, Monograph 3. (Original work published in 1926). [aBH]
(1930/1959) The function of the perceptual system. *Psychological Issues* 1959, Monograph 371–394. (Original work published 1930). [aBH]
Heister, G., Ehrenstein, W. H. & Schroeder-Heister, P. (1986) Spatial S-R compatibility effects with unimanual two-finger choice reactions for prone and supine hand positions. *Perception and Psychophysics* 40:271–78. [aBH]
(1987) Spatial S-R compatibility with unimanual two-finger choice reactions: Effects of irrelevant stimulus location. *Perception and Psychophysics* 42:195–201. [aBH]
Helmholtz, H. von. (1909) *Handbuch der Physiologischen Optik,* (third edition). Voss. [aBH]
Henderson, J. M. (1994) Two representational systems in dynamic visual identification. *Journal of Experimental Psychology: General* 123:410–26. [aBH]
Henderson, J. M. & Anes, M. D. (1994) Roles of object-file review and type priming in visual identification within and across eye-fixations. *Journal of Experimental Psychology: Human Perception and Performance* 20:826–39. [aBH]
Heptulla-Chatterjee, S., Freyd, J. J. & Shffrar, M. (1996) Configural processing in the perception of apparent biological motion. *Journal of Experimental Performance: Human Perception and Performance* 22:916–29. [aBH]
Hershberger, W. A., ed. (1989) *Volitional actions: Conation and control.* North-Holland. [aBH]
Heuer, H. (1993) Structural constraints on bimanual movements. *Psychological Research/Psychologische Forschung* 55(2):83–98. [JD]
Hick, W. E. (1952) On the rate of gain of information. *Quarterly Journal of Experimental Psychology* 4:11–26. [aBH, RWP]



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

Hinde, R. A. (1966) *Animal behaviour: A synthesis of ethology and comparative psychology.* McGraw-Hill. [PC]
Hochberg, J. E. (1968) In the mind's eye. In: *Contemporary theory and research in visual perception,* ed. R. N. Haber. Holt, Rinehart & Winston. [JH]
(1970) Attention, organization and consciousness. In: *Attention: Contemporary theory and analysis,* ed. D. J. Mostofscky. Appleton-Century-Crofts. [JH]
(1978) *Perception, 2nd edition.* Prentice-Hall. [aBH]
(1997) The affordances of perceptual inquiry: Pictures are learned from the world, and what that fact might mean about perception quite generally. *The Psychology of Learning and Motivation* 36:15–44. [JH]
Hochberg, J. E. & Brooks, V. (1996) The perceptin of motion pictures. In: *Handbook of perception and cognition: Cognitive ecology,* ed. M. Friedman & E. Carterette. Academic Press. [JH]
Hoffman, J. (1993) *Vorhersage und Erkenntnis. Die Funktion von Antizipationen in der menschlichen Verhaltenssteuerung* [*Prediction and knowledge. The function of anticipation for human action control*]. Hogrefe. [aBH, WK]
Hoffman, J. & Sebald, A. (2000) Lernmechanismen zum Erwerb verhaltenssteuernden Wissens. [Learning mechanisms for the acquisition of action-controlling knowledge]. *Psychologische Rundschau* 51:1–9. [WK]
Hommel, B. (1993) Inverting the Simon effect by intention: Determinants of direction and extent of effects of irrelevant spatial information. *Psychological Research/Psychologische Forschung* 55:270–79. [aBH]
(1996a) The cognitive representation of action: Automatic integration of perceived action effects. *Psychological Research/Psychologische Forschung* 59:176–86. [aBH]
(1996b) S-R compatibility effects without response uncertainty. *Quarterly Journal of Experimental Psychology* 49A:546–71. [aBH]
(1997) Toward an action concept model of stimulus-response compatibility. In: *Theoretical issues in stimulus-response compatibility,* ed. B. Hommel & W. Prinz. North-Holland. [arBH]
(1998a) Observing one's own action – and what it leads to. In: *Systems theory and a priori aspects of perception,* ed. J. S. Jordan. North-Holland. [arBH]
(1998b) Event files: Evidence for automatic integration of stimulus-response episodes. *Visual Cognition* 5:183–216. [aBH, JI, RWP]
(1998c) Automatic stimulus-response translation in dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance* 24:1368–84. [rBH]
(2001a) The prepared reflex: Automaticity and control in stimulus-response
(2001b) Intentional control of automatic stimulus-response translation. In: *Interaction between dissociable conscious and nonconscious processes,* ed. Y Rossetti & A. Revonsuo. MIT Press. [aBH]
Hommel, B. & Eglau, B. (in press) Control of stimulus-response translation in dual-task performance. *Psychological Research.* [rBH]
Hommel, B., Gehrke, J. & Knuf, L. (2000a) Hierarchical coding in the perception and memory of spatial layouts. *Psychological Research* 64:1–10. [rBH]
Hommel, B. & Knuf, L. (2000) Action related determinants of spatial coding in perception and memory. In: *Spatial cognition II: Integrating abstract theories, empirical studies, formal methods, and practical applications,* ed. C. Freska, W. Brauer, C. Habel & K. F. Wender. Springer. [rBH]
Hommel, B., Knuf, L. & Gehrke, J. (2001) Action-induced cognitive organization of spatial maps. (submitted). [rBH]
Hommel, B., Pösse, B. & Waszak, F. (2000b) Contextualization in perception and action. *Psychologica Belgica* 40:227–45. [rBH, WK]
Hommel, B. & Prinz, W., eds. (1997) *Theoretical issues in stimulus-response compatibility.* North-Holland. [aBH]
Hommel, B. & Schneider, W. X. (in press) Visual attention and the selection of manual responses: Distinct mechanisms operating on the same codes. *Visual Cognition.* [JD, arBH, WK]
Houghton, G. & Hartley, T. (1995) Parallel models of serial behavior: Lashley revisited. *PSYCHE* 2(25). http://psyche.cs.monash.edu.au/ [JB]
Howard, I. P. (1982) *Human visual orientation.* Wiley. [aBH]
Howard, I. P. & Templeton, W. B. (1966) *Human spatial orientation.* Wiley. [aBH]
Hull, C. L. (1952) *A behavior system: An introduction to behavior theory concerning the individual organism.* Yale University Press. [aBH]
Hyman, R. (1953) Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology* 45:188–96. [RWP]
Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, , J. C. & Rizzolatti, G. (1999) Cortical mechanisms of human imitation. *Science* 286(5449):2526–28. [TC]
Imamizu, H., Miyauchi, S., Tamada, T., Sasaki, Y., Takino, R., Putz, B., Yoshioko, T. & Kawato, M. (2000) Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature* 403(6766):192–95. [LP]
James, W. (1890/1981) *The principles of psychology.* Macmillan/Harvard University Press. (Original work published 1890). [arBH, WK]
Jeannerod, M. (1983) How do we direct our actions in space? *In: Spatially oriented behavior,* ed. A. Hein & M. Jeannerod. Springer. [aBH]
(1984) The contribution of open-loop and closed-loop control modes in prehension movements. In: *Prepatory states and processes: Proceedings of the Franco-American conference,* ed. S. Kornblum & J. Requin. Erlbaum. [rBH]
(1988) *The neural and behavioural organization of goal-directed movements.* Oxford University Press. [aBH]
(1990) The representation of the goal of an action and its role in the control of goal-directed movements. In: *Computational neuroscience,* ed. E. L. Schwartz. MIT Press. [aBH]
(1997) *The cognitive neuroscience of action.* Blackwell. [aBH, SV]
(1999) To act or not to act: Perspectives on the representation of actions (The 25th Bartlett Lecture). *Quarterly Journal of Experimental Psychology* 52:1–29. [aBH]
(2001) Neural simulation of action: A unifying mechanism for motor cognition. *Neuroimage* 14(1):S103-S109. [TC]
Jeannerod, M., Arbib, M. A., Rizzolatti, G. & Sakata, H. (1995) Grasping objects: The cortical mechanisms of visuomotor transformation. *Trends in Neurosciences* 18:314–20. [SV]
Jellema, T. & Perrett, D. I. (2002) Coding of visible and hidden actions. In: *Attention and performance, vol. XIX: Common mechanisms in perception and action,* ed. W. Prinz & B. Hommel. Oxford University Press. [rBH]
Johansson, G. (1950) *Configurations in event perception.* Almquist and Wiksells. [aBH]
(1973) Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics* 14:201–11. [aBH]
Jordan J. S., Kerzel, D., Stork, S., Knuf, L. & Müsseler, J. (2002) Action planning affects spatial localization. In: *Attention and performance XIX: Common mechanisms in perception and action,* ed. W. Prinz & B. Hommel. Oxford University Press. [JSJ]
Jürgens, M. & Dinse, H. R. (1997a) Use-dependent plasticity of SI cortical hindpaw neurons induced by modification of walking in adult rats: A model for age related alterations. *Society for Neuroscience Abstracts* 23:1800. [HRD]
(1997b) Differential effects of the $Ca^{2+}$-influxblocker nimodipine on receptive field properties and response latencies of somatosensory cortical neurons in aged rats. *Institut für Neuroinformatik, Ruhr-Universität Bochum, Internal Report* 97–10:1–23. [HRD]
Jusczyk, P. (1986) Speech perception. In: *Handbook of perception and human performance. Vol. 2: Cognitive processes and performance,* ed. K. R. Boff, L. Kaufman & J. P. Thomas. Wiley. [BG, aBH]
Kadar, E. & Shaw, R. E. (2000) Toward an ecological field theory of perceptual control of information. *Ecological Psychology* 12(2):141–80. [RES]
Kaelbling, L. P., Littman, M. L. & Moore, A. W. (1996) Reinforcement learning: A survey. *Journal of Artifical Intelligence Research* 4:237–85. [EC]
Kahneman, D. (1973) *Attention and effort.* Prentice Hall. [aBH]
Kahneman, D. & Treisman, A. M. (1984) Changing view of attention and automaticity. In: *Varieties of attention,* ed. R. Parasuraman, R. Davies & J. Beatty. Academic Press. [aBH]
Kahneman, D., Treisman, A. & Gibbs, B. (1992) The reviewing of object files: Object-specific integration of information. *Cognitive Psychology* 24:175–219. [aBH, JI, RDO]
Kalaska, J. F., Scott, S. H., Cisek, P. & Sergio, L. E. (1997) Cortical control of reaching movements. *Current Opinion in Neurobiology* 7:849–59. [PC]
Kalaska, J. F., Sergio, L. E. & Cisek, P. (1998) Cortical control of whole-arm motor tasks. In: *Sensory guidance of movement, Novartis Foundation Symposium No. 218,* ed. M. Glickstein. Wiley. [PC]
Kandel, S., Orliaguet, J.-P. & Boë, L.-J. (1995) Visual perception of motor anticipation in handwriting: Influence of letter size and movement velocity. In: *Studies on perceptiona and action, vol. III,* ed. B. G. Bardy, R. J. Bootsma & Y. Guiard. Erlbaum. [aBH]
(2000) Detecting anticipatory events in handwriting movements. *Perception* 29(8):953–64. [TC]
Kaplan, S., Weaver, M. & French, R. (1990) Active symbols and internal models: Towards a cognitive connectionism. *AI and Society* 4:51–71. [EC]
Katz, D. (1960) Sozialpsychologie. In: *Handbuch der Psychologie,* 2nd edition, ed. D. Katz & R. Katz. Schwabe. [aBH]
Keele, S. W. (1968) Movement control in skilled motor performance. *Psychological Bulletin* 70:387–403. [aBH]
Keele, S. W., Cohen, A. & Ivry, R. (1990) Motor programs: Concepts and issues. In: *Motor representation and control. Attention and performance XIII,* ed. M. Jeannerod. Erlbaum. [aBH]
Kelso, J. A. S. (1995) *Dynamic patterns: The self organization of brain as behavior.* MIT Press. [RES]
Kelso, J. A. S. & Kay, B. A. (1987) Information and control: A macroscopic analysis of perception-action coupling. In: *Perspectives on perception and action,* ed. H. Heuer & A. F. Sanders. Erlbaum. [rBH]
Kerzel, D., Jordan, J. S. & Müsseler, J. (2001) The role of perceptual anticipation in the localization of the final position of a moving target. *Journal of Experimental Psychology: Human Perception and Performance* 27(4):829–40. [JSJ]


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

Kim, J.-N. & Shadlen, M. N. (1999) Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nature Neuroscience* 2:176–85. [PC]

Kim, N.-G. & Effken, J. A. (1995) How the future is specified: A hypothesis. In: *Studies in perception and action III,* ed. B. G. Bardy, R. J. Bootsma & Y. Guiard. Erlbaum. [N-GK]

Kim, N.-G., Effken, J. A. & Carello, C. (1998) Perceiving the severity of contacts between two objects. *Ecological Psychology* 10:93–127. [N-GK]

Kim, N.-G., Effken, J. A. & Shaw, R. (1995) Perceiving persistence under change and over structure. *Ecological Psychology* 7:217–56. [N-GK]

Kim, N.-G. & Turvey, M. T. (1998) Optical flow fields and Bernstein's "modeling of the future." In: *Progress in motor control,* ed. M. L. Latash. Human Kinetics. [N-GK]

Kim, N.-G., Turvey, M. T. & Carello, C. (1993) Optical information about the severity of upcoming contacts. *Journal of Experimental Psychology: Human Perception and Performance* 19:179–93. [N-GK]

Kinsbourne, M. (1982) Hemispheric specialization and the growth of human understanding. *American Psychologist* 37(4):411–20. [EC]

Klatzky, R. L., Pellegrino, J., McCloskey, B. P. & Doherty, S. (1989) Can you squeeze a tomato? The role of motor representations in semantic sensibility judgments. *Journal of Memory and Language* 28:56–77. [aBH]

Klein, R. M. (1980) Does oculomotor readiness mediate cognitive control of visual attention? In: *Attention and performance, vol. VIII,* ed. R. S. Nickerson. Erlbaum. [arBH, JI]

Klein, R. M. & Pontefract, A. (1994) Does oculomotor readiness mediate cognitive control of visual attention? Revisited! In: *Attention and performance XV: Conscious and nonconscious information processing,* ed. C. Umiltà & M. Moscovitch. MIT Press. [rBH, JI]

Klotz, W. & Neumann, O. (1999) Motor activation without conscious discrimination in metacontrast masking. *Journal of Experimental Psychology: Human Perception and Performance* 25:976–92. [rBH]

Knuf, L. (1997) Ideomotorische Phänomene: Neue Fakten für ein altes Problem. Entwicklung eines Paradigmas zur kinematischen Analyse induzierter Mitbewegungen. Unpublished doctoral dissertation, University of Munich, Germany. [aBH]

Knuf, L., Aschersleben, G. & Prinz, W. (2000) An analysis of ideomotor action. *Journal of Experimental Psychology: General* 130:779–98. [aBH]

Koenderink, J. J. (1990) The brain, a geometry engine. Special Issue: Domains of mental functioning: Attempts at a synthesis. *Psychological Research* 52:122–27. [aBH]

Kohler, I. (1964) The formation and transformation of the perceptual world. *Psychological Issues* 3:1–173. [aBH]

Köhler, W. (1929) *Gestalt psychology.* Liveright. [aBH]

Konorski, J. (1967) *Integrative activity of the brain.* Chicago University Press. [aBH]

Kornblum, S., Hasbroucq, T. & Osman, A. (1990) Dimensional overlap: Cognitive basis for stimulus-response compatibility – A model and taxonomy. *Psychological Review* 97:253–70. [arBH, RWP, AFS]

Kortenkamp, D., Bonasso, R. P. & Murphy, R., eds. (1998) *Artificial Intelligence and mobile robots: Case studies of successful robot systems.* MIT Press. [JB]

Kowler, E. (1990) *Eye movements and their role in visual and cognitive processes.* Elsevier. [aBH]

Kringelbach, M. & Engell, A. (1997) Development and plasticity in the brain. *Psyke and Logos* 18:266–86. [MOB]

Kritikos, A. (in preparation) Perceptual processing and the modulation of distractor interference in reach-to-grasp paradigms. [LP]

Kugler, P. N. & Turvey, M. T. (1987) *Information, natural law, and the self-assembly of rhythmic movements.* Erlbaum. [MJR, RES]

Kunde, W. (2001) Response-effect compatibility in manual choice reaction tasks. *Journal of Experimental Psychology: Human Perception and Performance* 27:387–94. [WK]

Kunde, W., Hoffmann, J. & Zellmann, P. (in press) The impact of anticipated action effects on action planning. *Acta Psychologica.* [WK]

LaBerge, D. & Samuels, S. J. (1974) Toward a theory of automatic information processing in reading. *Cognitive Psychology* 6:293–323. [JI]

Lackner, J. R. & DiZio, P. (1994) Rapid adaptation to Coriolis force perturbations of arm trajectory. *Journal of Neurophysiology* 72:299–313. [PC]

Lakoff, G. (1999) *Philosophy in the flesh.* Basic Books. [DCR]

Lamme, V. (2000) Neural mechanisms of visual awareness: A linking proposition. *Brain and Mind* 1:385–406. [MOB]

Lane, P. C. R., Cheng, P. C.-H. & Gobet, F. (2000) CHREST+: Investigating how humans learn to solve problems using diagrams. *Artificial Intelligence and the Simulation of Behaviour (AISB) Quarterly* 103:24–30. [PCRL]

(2001) Learning perceptual chunks for problem decomposition. In: *Proceedings of the Twenty-third Annual Meeting of the Cognitive Science Society, Edinburgh, UK,* pp. 528–33, ed. J. D. Moore & K. Stenning. Erlbaum. [PCRL]

Lashley, K. S. (1951) The problem of serial order in behavior. In: *Cerebral mechanisms in behavior,* ed. L. A. Jeffress. Wiley. [JB]

Lee, D. N. (1976) A theory of visual control of braking based on information about time-to-collision. *Perception* 5:437–59. [aBH, N-GK]

Lee, D. N., Lishman, J. R. & Thomson, J. A. (1982) Regulation of gait in long jumping. *Journal of Experimental Psychology: Human Perception and Performance* 8:448–59. [N-GK]

Leslie, A. M., Xu, F., Tremoulet, P. D. & Sholl, B. J. (1998) Indexing and the object concept: Developing "what" and "where" systems. *Trends in Cognitive Sciences* 2:10–18. [aBH]

Levelt, W. J. M. (1989) *Speaking. From intention to articulation.* MIT Press. [BG, RJH, aBH]

Lewin, K. (1926) Vorsatz, Wille und Bedürfnis. *Psychologische Forschung* 7:330–85. [aBH]

Liberman, A. M. (1957) Some results of research on speech perception. *Journal of the Acoustical Society of America* 29(1):117–23. [BG]

(1982) On finding that speech is special. *American Psychologist* 37:148–67. [aBH]

(1996) *Speech: A special code.* MIT Press. [BG]

Liberman, A., Cooper, F. S., Shankweiler, D. P. & Studdert-Kennedy, M. (1967) Perception of the speech code. *Psychological Review* 74:431–61. [BG, JH]

Liberman, A. M. & Mattingly, I. G. (1985) The motor theory of speech perception revisited. *Perception* 21:1–36. [rBH]

(1989) A specialization for speech-perception. *Science* 243(4890):489–94. [BG]

Lindenberger, U., Marsiske, M. & Baltes, P. B. (2000) Memorizing while walking: Increase in dual-task costs from young adulthood to old age. *Psychology and Aging* 15:417–36. [HRD]

Livesey, P. J. (1986) *Learning and emotion: A biological synthesis. Evolutionary processes, vol. 1.* Erlbaum. [JB]

Logan, G. D. (1984) On the ability to inhibit simple and choice reaction time responses: A model and a method. *Journal of Experimental Psychology: Human Perception and Performance* 10:276–91. [rBH]

Logan, G. D. & Gordon, R. D. (2001) Executive control of visual attention in dual task situations. *Psychological Review* 108:393–434. [NM]

Lonstein, J. S. & Stern, J. M. (1997) Role of the midbrain periaqueductal gray in maternal nurturance and aggression: *C-fos* and electrolytic lesion studies in lactating rats. *Journal of Neuroscience* 17(9):3364–78. [JB]

Lotze, R. H. (1852) *Medicinische Psychologie oder Physiologie der Seele.* Weidmann. [arBH, AFS]

Lupiañez, J. & Milliken, B. (1999) Inhibition of return and the attentional set for integrating vs. differentiating information. *Journal of General Psychology* 126:392–418. [rBH]

MacCorquodale, K. & Meehl, P. E. (1954) Edward C. Tolman. In: *Modern learning theory: A critical analysis of five examples,* ed. W. K. Estes, S. Koch, K. MacCorquodale, P. E. Meehl, C. G. Mueller, Jr., W. N. Schoenfeld & W. S. Verplanck. Appleton-Century-Crofts. [EC]

Macfarlane, D. A. (1930) The role of kinesthesis in maze learning. *University of California Publications in Psychology* 4:277–305. [rBH]

Mackay, D. (1987) *The organization of perception and action: A theory of language and other cognitive skills.* Springer. [RJH, aBH]

Mackay, D., Allport, A., Prinz, W. & Scheerer, E. (1987) Relationships and modules within language perception and production: An introduction. In: *Language perception and production: Relationships between listening, speaking, reading, and writing,* ed. A. Allport, D. MacKay, W. Prinz & E. Scheerer. Academic Press. [aBH]

MacKay, D. & Mittelstaedt, H. (1974) Visual stability and motor control (reafference revisited). In: *Cybernetics and bionics,* ed. D. Keidel et al. Oldenbourg. [aBH]

Mark, L. S. (1987) Eye-height-scaled information about affordances: A study of sitting and climbing. *Journal of Experimental Psychology: Human Perception and Performance* 13:361–70. [MJR]

Marr, D. C. (1969) A theory of cerebellar cortex. *Journal of Physiology* 202:437–70. [PC]

Marr, D. C. (1982) *Vision.* W. F. Freeman. [N-GK, AFS]

Mashhour, M. (1964) *Psychophysical relations in the perception of velocity.* Almquist. [aBH]

Massad, C. M., Hubbard, M. & Newston, D. (1979) Selective perception for events. *Journal of Experimental Social Psychology* 15:513–32. [aBH]

Massaro, D. (1990) An information-processing analysis of perception and action. In: *Relationships between perception and action: Current approaches,* ed. O. Neumann & W. Prinz. Springer. [aBH]

Mates, J. & Aschersleben, G. (2000) Sensorimotor synchronization: The influence of temporally displaced auditory feedback. *Acta Psychologica* 104:29–44. [aBH]

Mates, J., Radil, T. & Pöppel, E. (1992) Cooperative tapping: Time control under different feedback conditions. *Perception and Psychophysics* 52:691–704. [aBH]

Matin, L. (1982) Visual localization and eye movements. In: *Symposium on the study of motion perception,* ed. W. A. Wagenaar, A. H. Wertheim & H. W. Leibowitz. Plenum Press. [aBH]



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

May, C. P., Kane, M. J. & Hasher, L. (1995) Determinants of negative priming. *Psychological Bulletin* 118:35–54. [aBH]
McCann, R. S. & Johnston, J. C. (1992) Locus of the single-channel bottleneck in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance* 18:471–84. [aBH]
McClelland, J. L. (1979) On the time relations of mental processes: An examination of systems of processes in cascade. *Psychological Review* 86:287–330. [AFS]
McKoon, G., Ratcliff, R. & Seifert, C. (1989) Making the connection: Generalized knowledge structures in story understanding. *Journal of Memory and Language* 28:711–334. [JMZ]
Meijer, O. G. & Roth, K., eds. (1988) *The motor-action controversy.* North-Holland. [rBH]
Meiran, N. (2000a) Modeling cognitive control in task-switching. *Psychological Research* 63:234–49. [NM]
(2000b) Reconfiguration of stimulus task-sets and response task-sets during task switching. In: *Control of cognitive processes: Attention and performance XVIII,* ed. S. Monsell & J. Driver. MIT Press. [arBH]
Meiran, N., Adi-Japha, E. & Kessler, Y. (in preparation) A general model-based analytic framework for task switching in speeded classification tasks. [NM]
Meiran, N., Chorev, Z. & Sapir, A. (2000) Component processes in task switching. *Cognitive Psychology* 41:211–53. [rBH, NM]
Meltzoff, A. N. (1995) Understanding the intentions of others: Re-enactment of intended act by 18-months-old children. *Developmental Psychology* 31:838–50. [aBH]
Meltzoff, A. N., Kuhl, P. K. & Moore, M. K. (1991) Perception, representation, and the control of action in newborns and young infants: Towards a new synthesis. In: *Newborn attention: Biological constraints and the influence of experience,* ed. M. J. Weiss & P. R. Zelazo. Ablex Press. [aBH]
Meltzoff, A. N. & Moore, M. K. (1977) Imitation of facial and manual gestures by human neonates. *Science* 198:75–78. [aBH]
Meltzoff, A. N. & Prinz, W. (in press) *The imitative mind: Development, evolution, and brain bases.* Cambridge University Press. [aBH]
Merigan, W. H. & Maunsell, J. H. (1993) How parallel are the primate visual pathways? *Annual Review of Neuroscience* 16:369–402. [aBH]
Metzger, W. (1975) *Gesetze des Sehens.* Waldemar Kramer. [aBH]
Meyer, D. E. & Gordon, P. C. (1985) Speech production: Motor programming of phonetic features. *Journal of Memory and Language* 24:3–26. [aBH]
Meyer, D. E. & Kieras, D. E. (1997) A computational theory of cognitive processses and multiple task performance: Part 1, Basic mechanisms. *Psychological Review* 104:3–75. [aBH]
Miall, R. C. & Wolpert, D. M. (1996) Forward models for physiological motor control. *Neural Networks* 9:1265–79. [PC]
Michaels, C. F. (1988) S-R compatibility between response position and destination of apparent motion: Evidence of the detection of affordances. *Journal of Experimental Psychology: Human Perception and Performance* 14:231–40. [aBH]
Michaels, C. F. & Carello, C. (1981) *Direct perception.* Prentice Hall. [aBH, MJR]
Michaels, C. F. & Stins, J. F. (1997) An ecological approach to stimulus-response compatibility. In: *Theoretical issues in stimulus-response compatibility,* ed. B. Hommel & W. Prinz. North-Holland. [aBH, MJR]
Middleton, F. A. & Strick, P. L. (2000a) Basal ganglia and cerebellar loops: Motor and cognitive circuits. *Brain Research Reviews* 31:236–50. [PC]
(2000b) Basal ganglia output and cognition: Evidence from anatomical, behavioral, and clinical studies. *Brain and Cognition* 42(2):183–200. [JB]
Miller, G. A., Galanter, E. & Pribram, K. H. (1960) *Plans and the structure of behavior.* Holt, Reinhart & Winston. [aBH, JI, WK]
Miller, J. (1982) Discrete versus continuous stage models of human information processing: In search of partial output. *Journal of Experimental Psychology: Human Perception and Performance* 8:273–96. [AFS]
(1988) Discrete and continuous models of human information processing: In search of partial output. *Acta Psychologica* 67:191–257. [AFS]
Miller, J., van der Ham, F. & Sanders, A. F. (1995) Overlapping stage models and the additive factor method. Acta Psychologica 90:11–28. [AFS]
Miller, N. E. & Dollard, J. (1941) *Social learning and imitation.* Yale University Press. [aBH]
Millikan, R. G. (1984) *Language, thought, and other biological categories.* MIT Press. [RGM]
(1993) *White Queen psychology and other essays for Alice.* MIT Press. [RGM]
(1996) Pushmi-pullyu representations. In: *Philosophical perspectives,* vol. IX, ed. J. Tomberlin. Ridgeview. [RGM]
(forthcoming) *Varieties of meaning: The Jean Nicod Lectures.* MIT Press. [RGM]
Milner, A. D. & Dijkerman, H. C. (in press) Direct and indirect visual routes to action. In: *Varieties of unconscious processing: New findings and new comparisons,* ed. B. de Gelder, E. H. F. de Haan & C. A. Heywood. Oxford University Press. [LP]
Milner, A. D. & Goodale, M. A. (1995) *The visual brain in action.* Oxford University Press. [arBH, LP, SV, DAW]
Mink, J. W. (1996) The basal ganglia: Focused selection and inhibition of competing motor programs. *Progress in Neurobiology* 50(4):381–425. [JB]
Mischel, W. (1996) From good intentions to willpower. In: *The psychology of action: Linking cognition and motivation to behavior,* ed. P. M. Gollwitzer & J. A. Bargh. Guilford Press. [aBH]
Monsell, S. & Driver, J., eds. (2000) *Control of cognitive processes. Attention and performance XVIII.* MIT Press. [aBH, NM]
Morin, R. E. & Grant, D. A. (1955) Learning and performance on a keypressing task as a function of the degree of spatial stimulus-response correspondence. *Journal of Experimental Psychology* 49:39–47. [aBH]
Morris, M. W. & Murphy, G. L. (1990) Converging operations on a basic level in event taxonomies. *Memory and Cognition* 18:407–18. [JMZ]
Movellan, J. R. & McClelland, J. L. (2001) The Morton-Massaro law of information integration. *Psychological Review* 108:103–48. [MOB]
Münsterberg, H. (1888) *Beiträge zur Experimentellen Psychologie. Heft 1: Über willkürliche und unwillkürliche Vorstellungsverbindungen.* Mohr. [aBH]
Murray, J. D. (1990) *Mathematical biology.* Springer Verlag. [RES]
Müsseler, J. (1995) *Wahrnehmung und Handlungssteuerung. Effekte kompatibler und inkompatibler Reize bei der Initiierung und Ausführung von Reaktionssequenzen* [Perception and action planning. Effects of compatible and incompatible stimuli on initiation and execution of reaction sequences.] Shaker. [aBH]
(1999) How independent from action control is perception? An event-coding account for more equally-ranked crosstalks. In: *Cognitive contributions to the perception of spatial and temporal events,* ed. G. Aschersleben, T. Bachmann & J. Müsseler. Elsevier. [arBH]
Müsseler, J. & Hommel, B. (1997a) Blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human Perception and Performance* 23(3):861–72. [JD, arBH, CO]
(1997b) Detecting and identifying response-compatible stimuli. *Psychonomic Bulletin and Review* 4:125–29. [arBH]
Müsseler, J. & Prinz, W. (1996) Action planning during the presentation of stimulus sequences: Effects of compatible and incompatible stimuli. *Psychological Research* 59:48–63. [aBH]
Müsseler, J., Steininger, S. & Wühr, P. (2001) Can actions really affect perceptual processing? *The Quarterly Journal of Experimental Psychology. Section A: Human Experimental Psychology* 54:137–54. [aBH, WK]
Müsseler, J., Van der Heijden, A. H. C., Mahmud, S. H., Deubel, H. & Ertsey, S. (1999) Relative mislocalizations of briefly presented stimuli in the retinal periphery. *Perception and Psychophysics* 61:1646–61. [aBH]
Müsseler, J., Wühr, P. & Prinz, W. (2000) Varying the response code in the blindness to response-compatible stimuli. *Visual Cognition* 7:743–67. [aBH]
Murthy, V. N. & Fetz, E. E. (1992) Coherent 25- to 35 Hz oscillations in the sensori-motor cortex of awake behaving monkeys. *Proceedings of the National Academy of Science USA* 89:5670–74. [aBH]
Nadel, J. & Butterworth, G., eds. (1999) *Imitation in infancy.* Cambridge University Press. [aBH]
Nattkemper, D. & Prinz, W. (2001) Impact of task demands on spatial stimulus-response compatibility. (submitted). [aBH]
Navon, D. & Gopher, D. (1979) On the economy of the human processing system. *Psychological Review* 86:214–25. [aBH, AFS]
Neafsey, E. J., Bold, E. L., Haas, G., Hurley-Gius, K. M., Quirk, G., Sievert, C. E. & Terreberry, R. R. (1986) The organization of the rat motor cortex: A microstimulation mapping study. *Brain Research* 396:77–96. [HRD]
Neill, W. T. (1997) Episodic retrieval in negative priming and repetition priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23:1291–305. [aBH]
Neisser, U. (1963) Decision-time without reaction-time: Experiments in visual scanning. *American Journal of Psychology* 76:376–85. [aBH]
(1967) *Cognitive psychology.* Appleton-Century-Crofts. [JH, rBH]
(1985) The role of invariant structures in the control of movement. In: *Goal-directed behavior: The concept of action in psychology,* ed. M. Frese & J. Sabini. Erlbaum. [aBH]
Nelson, J. I., Salin, P. A., Munk, M. H. J., Arzi, M. & Bullier, J. (1992) Spatial and temporal coherence in cortico-cortical connections: A cross-correlation study in areas 17 and 18 in the cat. *Visual Neuroscience* 9:21–37. [aBH]
Neumann, O. (1987) Beyond capacity: A functional view of attention. In: *Perspectives on perception and action,* ed. H. Heuer & A. F. Sanders. Erlbaum. [aBH]
(1990) Visual attention and action. In: *Relationships between perception and action: Current approaches,* ed. O. Neumann & W. Prinz. Springer. [aBH]
(1996) Theories of attention. In: *Handbook of perception and action. Vol. 3: Attention,* ed. O. Neumann & A. F. Sanders. Academic Press. [aBH]
Newell, A. & Simon, H. A. (1972) *Human problem solving.* Prentice-Hall. [JMZ]
Nickerson, R. S. (1972) Binary-classification reaction time: A review of some


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

studies of human information-processing capabilities. *Psychonomic Monograph Supplements* 4:275–88. [aBH]
Norman, D. A. (1969) *Memory and attention: An introduction to human information processing.* Wiley. [rBH, JI]
Norman, D. A. & Shallice, T. (1986) Attention to action: Willed and automatic control of behaviour. In: *Consciousness and self regulation,* ed. R. J. Davidson, G. E. Schwartz & D. Shapiro. Plenum. [MOB]
Oades, R. D., Zimmermann, B. & Eggers, C. (1996) Conditioned blocking in patients with paranoid, non-paranoid psychosis or obsessive-compulsive disorder: Association with symptoms, personality and monoamine metabolism. Journal of Psychiatric Research 30:369–90. [RDO]
Ohlsson, C. R. & Gettner, S. N. (1995) Object-centered direction selectivity in the macaque supplementary eye-field. *Science* 269:985–88. [aBH]
Olivetti Belardinelli, M. (1974/1986) *La costruzione della realtà come problema psicologico.* Boringhieri. (3rd enlarged edition, 1986). [MOB]
(1981–82) Lernen und motiviertes Verhalten: Ein Kognitiv-systhemisches Modell. *Verhalten* 3:95–120. [MOB]
Olivetti Belardinelli, M., Del Gratta, C., Di Matteo, R., Ferretti, A. & Romani, G. (2001) Intermodal analysis of sensory image generation by means of fMRI. Paper presented at the 8th EWIC, Saint-Malo, April 1–3, 2001. Submitted to *Neuroimage.* [MOB]
O'Regan, J. K. & Noë, A. (2001) A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences* 24(5). [JH, N-GK]
Paillard, J. (1949) Quelques données psychophysiologiques au déclenchement de la commande motrice [Some psychophysiological data relating to the triggering of motor commands]. *L'Année Psychologique* 48:28–47. [aBH]
Paillard, J., ed. (1991) *Brain and space.* Oxford University Press. [aBH]
Palmer, S. E. (1982) Symmetry, transformation, and the structure of perceptual systems. In: *Organization and representation in perception,* ed. J. Beck. Erlbaum. [aBH]
Park, J. & Kanwisher, N. (1994) Negative priming for spatial locations: Identity mismatching, not distractor inhibition. *Journal of Experimental Psychology: Human Perception and Performance* 20:613–23. [aBH]
Pashler, H. (1989) Dissociations and dependencies between speed and accuracy: Evidence for a two-component theory of divided attention in simple tasks. *Cognitive Psychology* 21:469–514. [aBH]
(1994) Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin* 116:220–44. [rBH]
Passingham, R. (1993) *The frontal lobes and voluntary action.* Oxford University Press. [aBH]
Payne, B R., Lomber, S. G., Villa, A. E. & Bullier, J. (1996) Reversible deactivation of cerebral network components. *Trends in Neurosciences* 19:535–42. [MOB]
Perani, D., Fazio, F., Borghese, N. A., Tettamanti, M., Ferrari, S., Decety, J. & Gilardi, M. C. (2001) Different brain correlates for watching real and virtual hand actions. *Neuroimage* 14(3):749–58. [TC]
Perrett, D. I., Harris, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., Citty, A. J., Hietanen, J. K. & Ortega, J. E. (1989) Framework of analysis for the neural representation of animate objects and actions. *Journal of Experimental Biology* 146:87–113. [aBH]
Perrett, D. I., Hietanen, J. K., Oram, M. W. & Benson, P. J. (1992) Organisation and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London* 335:25–30. [JB]
Peters, E. R., Pickering, A. D., Kent, A., Glasper, A., Irani, M., David, A. S., Day, S. & Hemsley, D. R. (2000) The relationship between cognitive inhibition and psychotic symptoms. *Journal of Abnormal Psychology* 109:386–95. [RDO]
Phaf, H. & Wolters, G. (1997) A constructivist and connectionist view on conscious and nonconscious processes. *Philosophical Psychology* 10:287–307. [GW]
Piaget, J. (1946) *La formation du symbole chez l'enfant.* Delachaux & Niestlé. [aBH]
Pickering, M. J. & Branigan, H. P. (1998) The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language* 39:633–51. [RJH]
Pickering, M. J. & Garrod, S. C. (submitted) Towards a mechanistic psychology of dialogue. Unpublished manuscript. [RJH]
Pisella, L., Grea, H., Tilikete, C., Vighetto, A., Desmurget, M., Rode, G., Boisson, D. & Rossetti, Y. (2000) An automatic pilot for the hand in the posterior parietal cortex: Toward a reinterpretation of Optic Ataxia. *Nature Neuroscience* 3(7):629–741. [LP]
Pisella, L. & Rossetti, Y. (2000) Interaction between conscious identification and non-conscious sensorimotor processing: Temporal constraints. In: *Beyond dissociation: Interaction between dissociated implicit and explicit processing,* ed. Y. Rossetti & A. Revonsuo. Benjamins. [LP]
Pleger, B., Dinse, H. R., Ragert, P., Schwenkreis, P., Malin, J. P. & Tegenthoff, M. (2001) Shifts in cortical representations predict human discrimination improvement. *Proceedings of the National Academy of Sciences USA* 98:12255–60. [HRD]
Pomerantz, J. R. & Kubovy, M. (1981) Perceptual organization: An overview. In: *Perceptual organization,* ed. M. Kubovy & J. R. Pomerantz. Erlbaum. [aBH]
Posner, M. I. (1973) *Cognition: An introduction.* Scott Foresman. [EC]
(1978) *Chronometric explorations of mind.* Erlbaum. [arBH, RWP]
(1980) Orienting of attention. *Quarterly Journal of Experimental Psychology* 32:3–25. [aBH]
(1986) Empirical studies of prototypes. In: *Noun classes and categorization,* ed. C. Craig. John Benjamins. [EC]
Posner, M. I. & Cohen, Y. (1980) Attention and the control of movement. In: *Tutorials in motor behavior,* ed. G. E. Stelmach & J. Requin. North-Holland. [aBH]
(1984) Components of visual orienting. In: *Control of language processes. Attention and performance, vol. X,* ed. H. Bouma & D. G. Bouwhuis. Erlbaum. [aBH]
Posner, M. I. & Mitchell, R. F. (1967) Chronometric analysis of classification. *Psychological Review* 74:392–409. [aBH]
Posner, M. I., Nissen, M. J. & Klein, R. M. (1976) Visual dominance: An information-processing account of its origins and significance. *Psychological Review* 83:157–71. [aBH]
Postman, L. (1955) Association theory and perceptual learning. *Psychological Review* 62:438–46. [aBH]
Powers, W. T. (1973) Feedback: Beyond behaviorism. *Science* 179:351–56. [PC]
Prablanc, C. & Pélisson, D. (1990) Gaze saccade orienting and hand pointing are locked to their goal by quick internal loops. In: *Attention and performance XIII,* ed. M. Jeannerod. Erlbaum. [rBH]
Prescott, T. J., Gurney, K., Gonzalez, F. M. & Redgrave, P. (forthcoming) The evolution of action selection. In: *Towards the whole Iguana,* ed. D. McFarland & O. Holland. MIT Press. [JB]
Pribram, K. H. (1971) *Languages of the brain. Experimental paradoxes and principles in neuropsychology.* Prentice Hall. [aBH]
Prinz, W. (1983) *Wahrnehmung und Tätigkeitssteuerung.* Springer. [AFS]
(1984a) Attention and sensitivity in visual search. *Psychological Research* 45:355–66. [rBH]
(1984b) Modes of linkages between perception and action. In: *Cognition and motor processes,* ed. W. Prinz & A. F. Sanders. Springer. [aBH]
(1987) Ideo-motor action. In: *Perspectives on perception and action,* ed. H. Heuer & A. F. Sanders. Erlbaum. [arBH]
(1990) A common coding approach to perception and action. In: *Relationships between perception and action: Current approaches,* ed. O. Neumann & W. Prinz. Springer. [aBH]
(1992) Why don't we perceive our brain states? *European Journal of Cognitive Psychology* 4:1–20. [arBH, SV]
(1997a) Perception and action planning. *European Journal of Cognitive Psychology* 9:129–54. [aBH]
(1997b) Why Donders has led us astray. In: *Theoretical issues in stimulus-response compatibility,* ed. B. Hommel & W. Prinz. North-Holland. [arBH]
(2002) An experimental approach to imitation. In: *The imitative mind: Development, evolution, and brain bases,* ed. A. M. Meltzoff & W. Prinz. Cambridge University Press (in press). [arBH]
Prinz, W., Aschersleben, G., Hommel, B. & Vogt, S. (1995) Handlungen als Ereignisse. In: *Das Gedächtnis, Probleme, Trends, Perspektiven,* ed. D. Dörner & E. v. d. Meer. Hogrefe. [aBH]
Proctor, R. W. & Reeve, T. G. (1986) Salient-feature coding operations in spatial precuing tasks. *Journal of Experimental Psychology: Human Perception and Performance* 12:277–85. [RWP]
Proctor, R. W. & Reeve, T. G., eds. (1990) *Stimulus-response compatibility.* Elsevier. [aBH]
Proctor, R. W., Reeve, T. G. & Weeks, D. J. (1990) A triphasic approach to the acquisition of response-selection skill. *The Psychology of Learning and Motivation* 26:207–40. [rBH]
Proctor, R. W., Van Zandt, T., Lu, C. H. & Weeks, D. J. (1993) Stimulus-response compatibility for moving stimuli: Perception of affordances or directional coding? *Journal of Experimental Psychology Human Perception and Performance* 19:81–91. [aBH]
Proctor, R. W. & Vu, K.-P. L. (in press a) Action selection. In: *Experimental psychology,* ed. A. F. Healy & R. W. Proctor. (Volume 3 of the *Comprehensive handbook of psychology,* Editor-in-Chief I B. Weiner). Wiley. [RWP]
(in press b) Eliminating magnifying and reversing spatial compatibility effects with mixed location-relevant and irrelevant trials. In: *Common mechanisms in perception and action,* ed. W. Prinz & B. Hommel. Oxford University Press. [RWP]
Purves, D. (1988) *Body and brain.* Harvard University Press. [RES]
Pylyshyn, Z. (2001) Visual indexes, preconceptual objects, and situated vision. *Cognition* 80(1–2):127–58. [DCR]
Quinn, J. T. & Sherwood, D. E. (1983) Time requirements of changes in program and parameter variables in rapid ongoing movements. *Journal of Motor Behavior* 15:163–78. [aBH]



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

Rachlin, H. C. (1966) Scaling subjective velocity, distance, and duration. *Perception and Psychophysics* 1:77–82. [aBH]
Radvansky, G. A. (1999) Memory retrieval and suppression: The inhibition of situation models. *Journal of Experimental Psychology: General* 128:563–79. [JMZ]
Rayner, K. (1984) Visual selection in reading, picture perception, and visual search. A tutorial review. In: *Control of language processes. Attention and performance, vol. X,* ed. H. Bouma & D. G. Bouwhuis. Erlbaum. [aBH]
Redding, G. & Wallace, B. (1997) *Adaptive spatial alignment.* Erlbaum. [aBH]
Redgrave, P., Prescott, T. J. & Gurney, K. (1999) The basal ganglia: A vertebrate solution to the selection problem? *Neuroscience* 89:1009–23. [JB]
Reed, E. S. (1982) An outline of a theory of action systems. *Journal of Motor Behavior* 14:98–134. [aBH]
(1993) The intention to use a specific affordance. In: *Development in context: Acting and thinking in specific environments,* ed. R. H. Wozniak & K. W. Fischer. Erlbaum. [aBH]
(1996) *Encountering the world: Toward an ecological psychology.* Oxford University Press. [aBH]
Reinke, H. & Dinse, H. R. (1996) Functional characterization of cutaneous mechanoreceptor properties in aged rats. *Neuroscience Letters* 216:171–74. [HRD]
Rescorla, R. A. (1991) Associative relations in instrumental learning: The Eighteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology* 43B:1–23. [aBH]
Restle, F. (1982) Coding theory as an integration of Gestalt Psychology and information processing theory. In: *Organization and representation in perception,* ed. J. Beck. Erlbaum. [aBH]
Richardson, D. C. & Spivey, M. J. (2000) Representation, space and Hollywood Squares. Looking at things that aren't there anymore. *Cognition* 76(3):269–95. [DCR]
Richardson, D. C., Spivey, M. J. & Cheung, J. (2001) Motor representations in memory and mental models: Embodiment in cognition. *Proceedings of the Twenty-third Annual Meeting of the Cognitive Science Society,* pp. 867–72. Erlbaum. [rBH, DCR]
Ridderinkhof, K. R., van der Molen, M. W. & Bashore, T. R. (1995) Limits on the application of the additive factors logic: Violations of stage robustness suggest a dual-process architecture to explain flanker effects on target processing. *Acta Psychologica* 90:29–48. [AFS]
Riehle, A., Gruen, S., Diesmann, M. & Aertsen, A. (1997) Spike synchronization and rate modulation differentially involved in motor cortical function. *Science* 278:1950–53. [aBH]
Rifkin, A. (1985) Evidence for a basic level in event taxonomies. *Memory and Cognition* 13:538–56. [JMZ]
Riggio, L., Gawryszewski, L. & Umiltà, C. (1986) What is crossed in crossed-hand effects? *Acta Psychologica* 62:89–100. [aBH]
Rizzolatti, G. & Craighero, L. (1998) Spatial attention: Mechanisms and theories. In: *Advances in psychological science, vol. 2: Biological and cognitive aspects,* ed. M. Sabourin, F. Craik & M. Robert. Psychology Press. [aBH]
Rizzolatti, G., Fadiga, L., Gallese, V. & Fogassi, L. (1996) Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3:131–41. [aBH, SV]
Rizzolatti, G., Fogassi, L. & Gallese, V. (1997) Parietal cortex: From sight to action. *Current Opinions in Neurobiology* 7:562–67. [aBH]
(2001) Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience* 2:661–670. [rBH]
Rizzolatti, G., Luppino, G. & Matelli, M. (1998) The organization of the motor cortical system: New concepts. *Electroencephalography and Clinical Neurophysiology* 106:283–96. [SV]
Rizzolatti, G., Riggio, L., Dascola, I. & Umiltà, C. (1987) Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia* 25(1A):31–40. *Special Issue: Selective visual attention.* [arBH]
Rizzolatti, G., Riggio, L. & Sheliga, B. M. (1994) Space and selective attention. In: *Attention and performance XV: Conscious and nonconscious information processing,* ed. C. Umiltà & M. Moscovitch. MIT Press. [aBH]
Roberts, S. (1987) Evidence for distinct serial processes in animals: The multiplicative factors method. *Animal Learning and Behavior* 15:135–73. [AFS]
Roelfsema, P. R., Engel, A. K., Koenig, P. & Singer, W. (1997) Visuomotor integration is associated with zero time-lag synchronization among cortical areas. *Nature* 385:157–61. [aBH, RDO]
Rogers, R. D. & Monsell, S. (1995) The cost of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General* 124:207–31. [NM]
Rosch, E. (1978) Principles of categorization. In: *Cognition and categorization,* ed. E. Rosch & B. Lloyd. Erlbaum. [EC]
Rosen, R. (1991) *Life itself.* Columbia University Press. [N-GK]
Rosenbaum, D. A. (1980) Human movement initiation: Specification of arm, direction, and extent. *Journal of Experimental Psychology: General* 109:444–74. [aBH]
(1987) Successive approximations to a model of human motor programming. *Psychology of Learning and Motivation* 21:153–82. [aBH]
(1991) *Human motor control.* Academic Press. [aBH]
Rosenbaum, D. A., Marchak, F., Barnes, H. J., Vaughan, J., Slotta, J. D. & Jorgensen, M. J. (1990) Constraints for action selection: Overhand versus underhand grips. In: *Attention and performance XIII: Motor representation and control,* ed. M. Jeannerod. Erlbaum. [rBH]
Rosenbaum, D. A., Meulenbroek, R. G., Vaughan, J. & Jansen, C. (2001) Posture-based motion planning: Applications to grasping. *Psychological Review* 108:709–34. [DAR]
Rosenbaum, D. A., Weber, R. J., Hazelett, W. M. & Hindorff, V. (1986) The parameter remapping effect in human performance: Evidence from tongue twisters and finger fumblers. *Journal of Memory and Language* 25:710–25. [arBH]
Rosenbloom, P. S. & Newell, A. (1987) An integrated computational model of stimulus-response compatibility and practice. *Psychology of Learning and Motivation* 21:3–49. [aBH]
Rosenblueth, A., Wiener, N. & Bigelow, J. (1943) Behavior, purpose and teleology. *Philosophy of Science* 10:18–24. [PC]
Rossetti, Y. & Pisella, L. (2002) Several "vision for action" systems: A guide to dissociating and integrating dorsal and ventral functions. In: *Common mechanisms in perception and action: Attention and performance, vol. XIX,* ed. W. Prinz & B. Hommel. Oxford University Press. [LP]
Rossetti, Y., Pisella, L. & Pelisson, D. (2000) Eye blindness and hand sight: Temporal constraints and visuo-motor network. *Visual Cognition* 7(6):785–809. [LP]
Rossetti, Y. & Revonsuo, A., eds. (in press) *Interaction between dissociable conscious and nonconscious processes.* John Benjamins. [aBH]
Rossetti, Y., Stelmach, G., Desmurget, M., Prablanc, C. & Jeannerod, M. (1994) Influence of viewing the static hand prior to movement onset on pointing kinematics and accuracy. *Experimental Brain Research* 101:323–30. [LP]
Rowe, J. B., Toni, I., Josephs, O., Frackowiak, R. S. & Passingham, R. E. (2000) The prefrontal cortex: Response selection or maintenance within working memory? *Science* 288:1656–60. [PC]
Ruby, P. & Decety, J. (2001) Effect of subjective perspective-taking during simulation of action: A PET investigation of agency. *Nature Neuroscience* 4(5):546–50. [TC]
Rumelhart, D. E. (1977) Understanding and summarizing brief stories. In: *Basic processes in reading: Perception and comprehension,* ed. D. Laberge & S. J. Samuels. Erlbaum. [JMZ]
Runeson, S. (1974) Constant velocity – not perceived as such. *Psychological Research* 37:3–23. [aBH]
Runeson, S. & Frykholm, G. (1981) Visual perception of lifted weights. *Journal of Experimental Psychology: Human Perception and Performance* 7:733–40. [aBH]
(1983) Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender, recognition, and deceptive intention. *Journal of Experimental Psychology* 112:585–615. [MJR]
Ruthruff, E., Miller, J. & Lachmann, T. (1995) Does mental rotation require central mechanisms? *Journal of Experimental Psychology: Human Perception and Performance* 21:552–70. [aBH]
Sakata, H., Taira, M., Murata, A. & Mine, S. (1995) Neural mechanisms of visual guidance of hand action in the parietal cortex of the monkey. *Cerebral Cortex* 5(5):429–38. [aBH, CO]
Salmoni, A. W., Schmidt, R. A. & Walter, C. B. (1984) Knowledge of results and motor learning: A review and critical appraisal. *Psychological Bulletin* 95:355–86. [aBH]
Sanders, A. F. (1967) Some aspects of reaction processes. *Acta Psychologica* 27:115–30. [AFS]
(1977) Structural and functional aspects of the reaction process. In: *Attention and performance, vol. 6,* ed. S. Dornic. Erlbaum. [AFS]
(1980) Stage analysis of reaction processes. In: *Tutorials in motor behavior,* ed. G. E. Stelmach & J. Requin. North-Holland. [arBH]
(1983) Towards a model of stress and human performance. *Acta Psychologica* 53:61–97. [aBH, RDO, AFS]
(1990) Issues and trends in the debate on discrete vs. continuous processing of information. *Acta Psychologica* 74:123–67. [aBH]
(1995) Discrete vs. continuous processing: The fate of an incompletely processed perceptual dimension. *Acta Psychologica* 90:211–27. [AFS]
(1997) A summary of resource theories from a behavioral perspective. *Biological Psychology* 45:5–18. [AFS]
(1998) *Elements of human performance. Reaction processes and attention in human skills.* Erlbaum. [aBH, AFS]
Scannell, J. W., Burns, G. A., Hilgetag, C. C., O'Neil, M. A. & Young, M. P. (1999) The connectionist organization of the cortico-thalamic system of the cat. *Cerebral Cortex* 9:277–99. [HRD]


https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

Schank, R. (1972) Conceptual dependency: A theory of natural language processing. *Cognitive Psychology* 3:552–631. [JMZ]
Schank, R. C. & Abelson, R. P. (1977) *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures.* Erlbaum. [JMZ]
Scheerer, E. (1984) Motor theories of cognitive structure: A historical review. In: *Cognition and motor processes,* ed. W. Prinz & A. F. Sanders. Springer. [arBH]
Schmidt, R. A. (1975) A schema theory of discrete motor skill learning. *Psychological Review* 82:225–60. [aBH]
(1988) *Motor control and learning, 2nd edition.* Human Kinetics Publishers. [aBH]
Schubö, A. (1998) *Interferenz von Wahrnehmung und Handlungssteuerung.* Shaker. [aBH]
Schubö, A., Aschersleben, G. & Prinz, W. (2001) Interactions between perception and action in a reaction task with overlapping S-R assignments. *Psychological Research* 65:145–57. [aBH, CO]
Schurman, T., Klein, H., Beneke, M. & Traber, J. (1987) Nimodipine and motor deficits in the aged rat. *Neuroscience Research Communications* 1:9–15. [HRD]
Schyns, P., Goldstone, R. L. & Thibaut, J. P. (1988) The development of features in object concepts. *Behavioral and Brain Sciences* 21:1–54. [GW]
Seifert, C. M., McKoon, G., Abelson, R. P. & Ratcliff, R. (1986) Memory connections between thematically similar episodes. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 12:220–31. [JMZ]
Semjen, A. & Garcia-Colera, A. (1986) Planning and timing of finger-tapping sequences with a stressed element. *Journal of Motor Behavior* 18:287–322. [rBH]
Shadmehr, R. & Mussa-Ivaldi, F. A. (1994) Adaptive representation of dynamics during learning of a motor task. *Journal of Neuroscience* 14:3208–24. [PC]
Shallice, T. (1982) Specific impairments in planning. *Philosophical Transactions of the Royal Society of London B* 298:199–209. [MOB]
(1988) *From neuropsychology to mental structure.* Cambridge University Press. [MOB]
Shaw, R. E. & Bransford, J. (1977) *Perceiving, acting, and knowing: Toward an ecological psychology.* Erlbaum. [MJR]
Shaw, R. E., Flascher, O. M. & Kadar, E. E. (1995) Dimensionless invariants for intentional systems: Measuring the fit of vehicular activities to environmental layout. In: *Global perspectives on the ecology of human-machine systems,* ed. J. Flach, P. Hancock, J. Caird & K. Vicente. Erlbaum. [RES]
Shaw, R. E., Kadar, E. E., Sim, M. M. & Repperger, D. W. (1992) The intentional spring: A strategy for modeling systems that learn to perform intentional acts. *Journal of Motor Behavior* 24:3–28. [RES]
Shaw, R. E. & Kinsella-Shaw, J. (1988) Ecological mechanics: A physical geometry for intentional constraints. *Human Movement Science* 7:155–200. [MJR]
Shaw, R. E. & Todd, J. (1980) Abstract machine theory and direct perception. *Behavioral and Brain Sciences* 3:400–401. [RES]
Shebilske, W. L. (1977) Visuomotor coordination in visual direction and position constancies. In: *Stability and constancy in visual perception: Mechanisms and processes,* ed. W. Epstein. Wiley. [aBH]
Sheliga, B. M., Riggio, L. & Rizzolatti, G. (1994) Orienting of attention and eye movements. *Experimental Brain Research* 98:507–22. [JI]
Shiffrar, M. (2001) Movement and event perception. In: *Handbook of perception,* ed. E. B. Goldstein. Blackwell. [TC]
Shiffrar, M. & Freyd, J. J. (1990) Apparent motion of the human body. *Psychological Science* 1:257–64. [TC, aBH]
(1993) Timing and apparent motion path choice with human photographs. *Psychological Science* 4:379–84. [TC]
Shiffrin, R. M. & Schneider, W. (1977) Controlled and automatic human information processing, II. *Psychological Review* 84:127–90. [aBH]
Simon, J. R., Hinrichs, J. V. & Craft, J. L. (1970) Auditory S-R compatibility. Reaction time as a function of ear-hand correspondence and ear-response-location correspondence. *Journal of Experimental Psychology* 86:97–102. [aBH]
Singer, W. (1994) The organization of sensory motor representations in the neocortex: A hypothesis based on temporal coding. In: *Conscious and nonconscious information processing: Attention and performance XV,* ed. C. Umiltà & M. Moscovitch. MIT Press. [arBH]
Smolensky, P. (1988) On the proper treatment of connectionism. *Behavioral and Brain Sciences* 11:1–74. [EC]
Spengler, F., Godde, B. & Dinse, H. R. (1995) Effects of aging on topographic organization of somatosensory cortex. *NeuroReport* 6:469–73. [HRD]
Sperry, R. W. (1952) Neurology and the mind-brain problem. *American Scientist* 40:291–312. [BG]
Spijkers, W., Heuer, H., Steglich, C. & Kleinsorge, T. (2000) Specification of movement amplitudes for the left and right hands: Evidence for transient parametric coupling from overlapping-task performance. *Journal of Experimental Psychology: Human Perception and Performance* 26(3):1091–105. [JD]
Spivey, M. J. & Geng, J. J. (2001) Oculomotor mechanisms activated by imagery and memory: Eye movements to absent objects. *Psychological Research* 65:235–41. [DCR]
Spivey, M. J., Tyler, M., Richardson, D. C. & Young, E. (2000) Eye movements during comprehension of spoken scene descriptions. *Proceedings of the Twenty-second Annual Meeting of the Cognitive Science Society.* Erlbaum. [DCR]
Stagner, R. (1977) Homeostasis, discrepancy, dissonance. *Motivation and Emotion* 1:103–38. [MOB]
Stanfield, R. A. & Zwaan, R. A. (2001) The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science* 12(2):153–56. [DCR]
Steininger, S. (1999) *Handeln und Wahrnehmen: Eine experimentelle Analyse einer Wahrnehmungsbeeinträchtigung bei simultan ausgeführten Handlungen.* [Acting and perceiving: An experimental analysis of impaired perception with concurrent action.] Doctoral thesis. Shaker. [WK]
Sternberg, S. (1967) Two operations in character recognition: Some evidence from reaction-time measurements. *Perception and Psychophysics* 2:45–53. [aBH]
(1969) The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica* 30:276–315. [arBH, AFS]
(1975) Memory's scanning: New findings and current controversies. *Quarterly Journal of Experimental Psychology* 27:1–42. [AFS]
Stevanovski, B., Oriet, C. & Jolicoeur, P. (2001a) Blinded by headlights. (submitted). [CO]
(2001b) Blinded by headlights. Paper presented at the 11th Annual Meeting of the Canadian Society for Brain, Behaviour, and Cognitive Science, Quebec City, Canada. [CO]
Stevens, J. A., Fonlupt, P., Shiffrar, M. & Decety, J. (2000) New aspects of motion perception: Selective neural encoding of apparent human movements. *NeuroReport* 11(1):109–15. [TC]
Stock, A. & Hoffman, J. (2002) Intentional fixation of behavioral learning or how R-E learning blocks S-R learning. *European Journal of Cognitive Psychology* 14(1):1–27. [WK]
Stoet, G. (1998) The role of feature integration in action planning. Unpublished doctoral dissertation, University of Munich, Germany. [aBH]
Stoet, G. & Hommel, B. (1999) Action planning and the temporal binding of response codes. *Journal of Experimental Psychology: Human Perception and Performance* 25:1625–40. [JD, arBH, WK, CO, DCR]
Stoffels, E. J., van der Molen, M. W. & Keuss, P. G. J. (1989) An additive factor analysis of the effect(s) of location cues associated with auditory stimuli on stages of information processing. *Acta Psychologica* 70:161–97. [rBH]
Stränger, J. & Hommel, B. (1996) The perception of action and movement. In: *Handbook of perception and action. Vol. 1: Perception,* ed. W. Prinz & B. Bridgeman. Academic Press. [arBH]
Stratton, G. M. (1897) Vision without inversion of the retinal image. *Psychological Review* 4:341–60, 463–81. [aBH]
Straube, E. R. & Oades, R. D. (1992) *Schizophrenia: Empirical research and findings.* Academic Press. [RDO]
Sturgill, D. S. & Ferraro, F. R. (1997) Predicting negative priming: Effects of stereotypy, cognitive failures and anxiety. *Personality and Individual Differences* 23:291–304. [RDO]
Stürmer, B. (1997) *Organisationsprinzipien an der Schnittstelle zwischen Wahrnehmung und Handlung: Kompatibilitätseffekte unter Verwendung dynamischer Reiz- und Reaktionseigenschaften.* [Principles of organization at the gap between perception and action: Compatibility effects with dynamic stimulus and response features.] Shaker. [aBH]
Stürmer, B., Aschersleben, G. & Prinz, W. (2000) Correspondence effects with manual gestures and postures: A study on imitation. *Journal of Experimental Psychology: Human Perception and Performance* 26:1746–59. [aBH]
Stuyven, E., Dejonckheere, P. & Ghyselinck, M. (1999) The influence of spatial compatibility on exogenous and endogenous saccades. In: *Proceedings of the Eleventh Conference of the European Society for Cognitive Psychology,* ed. A. Vandierendonck, M. Brysbaert & K. Van der Goten. Academia Press. [rBH]
Taira, M., Mine, S., Georgopoulos, A. P., Murata, A. & Sakata, H. (1990) Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Experimental Brain Research* 83:29–36. [aBH]
Tanji, J & Shima, K. (1994) Role for supplementary motor area cells in planning several movements ahead. *Nature* 371:413–16. [JB]
Taylor, J. L. & McCloskey, D. (1990) Triggering of preprogrammed movements as reactions to masked stimuli. *Journal of Neurophysiology* 63:439–46. [LP]
Teichner, W. H. & Krebs, M. J. (1974) Laws of visual choice reaction time. *Psychological Review* 83:29–36. [aBH]
Ten Hoopen, G., Akerboom, S. & Raaymakers, E. (1982) Vibrotactual choice reaction time, tactile receptor systems and ideomotor compatibility. *Acta Psychologica* 50:143–57. [aBH]
Thayer, J. F. & Lane, R. D. (2000) A model of neurovisceral integration in emotion



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

regulation and dysregulation. *Journal of Affective Disorders* 61:201–16. [MOB]
Theeuwes, J. (1993) Visual selective attention: A theoretical analysis. *Acta Psychologica* 83:93–154. [AFS]
Thelen, E., Schöner, G., Scheier, C. & Smith, L. B. (2001) The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences* 24(1):1–34; discussion 34–86. [HRD]
Thelen, E. & Smith, L. B. (1994) *A dynamic systems approach to the development of cognition and action.* MIT Press. [aBH]
Theios, J. (1975) The components of response latency in simple human information processing tasks. In: *Attention and performance V,* ed. P. M. A. Rabbitt & S. Dornic. Academic Press. [aBH]
Thompson, D. W. (1942) *On growth and form.* Dover. [RES]
Tipper, S. P., Howard, L. A. & Houghton, G. (1998) Action-based mechanisms of attention. *Philosophical Transactions of the Royal Society of London B* 353:1385–93. [LP]
Toates, F. (1998) The interaction of cognitive and stimulus-response processes in the control of behaviour. *Neuroscience and Biobehavioral Reviews* 22:59–83. [PC]
Tolman, E. C. (1932) *Purposive behavior in animals and man.* Appleton-Century-Crofts. [JH, rBH]
(1948) Cognitive maps in rats and men. *Psychological Review* 53:189–208. [EC, JH]
(1959) Principles of purposive behavior. In: *Psychology: A study of science, vol. 2,* ed. S. Koch. McGraw-Hill. [rBH]
Tolman, E. C., Hall, C. S. & Bretnall, E. P. (1932) A disproof of the law of effect and a substitution of the laws of emphasis, motivation and disruption. *Journal of Experimental Psychology* 15:601–14. [rBH]
Treisman, A. (1960) Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology* 12:242–48. [JI]
(1988) Features and objects: The Fourteenth Bartlett Memorial Lecture. *The Quarterly Journal of Experimental Psychology* 40A:201–37. [aBH]
(1992) Perceiving and re-perceiving objects. *American Psychologist* 47:862–75. [aBH]
(1996) The binding problem. *Current Opinion in Neurobiology* 6:171–78. [aBH]
Treisman, A. M. & Gelade, G. (1980) A feature-integration theory of attention. *Cognitive Psychology* 12:97–136. [aBH]
Tucker, M. & Ellis, R. (1998) On the relations between seen objects and components of potential actions. *Journal of Experimental Psychology: Human Perception and Performance* 24(3):830–46. [TC, rBH, DCR]
Turvey, M. T. (1977) Preliminaries to a theory of action with reference to vision. In: *Perceiving, acting, and knowing. Toward an ecological psychology,* ed. R. Shaw & J. Bransford. Erlbaum. [arBH, MJR]
(1992) Affordances and prospective control: An outline of the ontology. *Ecological Psychology* 4:173–87. [N-GK]
(1996) Dynamic touch. *American Psychologist* 51:1134–52. [MJR]
Turvey, M. T. & Fitzpatrick, P. (1993) Commentary: Development of perception-action systems and general principles of pattern formation. *Child Development* 64:1175–90. [RES]
Umiltà, C. & Nicoletti, R. (1990) Spatial stimulus-response compatibility. In: *Stimulus-response compatibility: An integrated perspective,* ed. R. W. Proctor & T. G. Reeve. North-Holland. [RWP]
Ungerleider, L. G. & Mishkin, M. (1982) Two cortical visual systems. In: *Analysis of visual behavior,* ed. D. Ingle, M. A. Goodale & R. J. W. Mansfield. MIT Press. [aBH]
Van der Goten, K., Caessens, B., Lammertyn, J., De Vooght, G. & Hommel, B. (2001) The functional basis of backward-compatibility effects: Selecting emotional actions primes the perception of emotional words. (submitted). [rBH]
Van der Heijden, A. H. C. (1992) *Selective attention in vision.* Routledge. [aBH, AFS]
Van der Heijden, A. H. C., Müsseler, J. & Bridgeman, B. (1999a) On the perception of position. In: *Cognitive conributions to the perception of spatial and temporal events,* ed. G. Aschersleben, T. Bachmann & J. Müsseler. Elsevier. [aBH]
Van der Heijden, A. H. C., Van der Geest, J. N., De Leeuw, F., Krikke, K. & Müsseler, J. (1999b) Sources of position-perception error for small isolated targets. *Psychological Research* 62:20–35. [aBH]
Van Donkelaar, P. & Franks, I. M. (1991) Preprogramming vs. on-line control in simple movement sequences. *Acta Psychologica* 77:1–19. [rBH]
Van Leeuwen, C. & Raffone, A. (2001) Coupled nonlinear maps as models of a perceptual pattern and memory trace dynamics. *Cognitive Processing* 2:67–116. [MOB]
Van Leeuwen, C., Steyvers, M. & Nooter, M. (1997) Stability and intermittency in large-scale coupled oscillator models for perceptual segmentation. *Journal of Mathematical Psychology* 41:1–26. [GW]
Van Selst, M. & Jolicoeur, P. (1997) Decision and response. *Cognitive Psychology* 33:266–307. [rBH]
Viviani, P. & Stucchi, N. (1989) The effect of movement velocity on form perception: Geometric illusions in dynamic displays. *Perception and Psychophysics* 46:266–74. [aBH]
Viviani, P. & Stucchi, N. (1992) Motor-perceptual interactions. In: *Tutorials in motor behavior, II,* ed. G. E. Stelmach & J. Requin. Elsevier. [aBH]
Viviani, P., Baud-Bovy, G. & Redolfi, M. (1997) Perceiving and tracking kinesthetic stimuli: Further evidence of motor-perceptual interactions. *Journal of Experimental Psychology: Human Perception and Performance* 23:1232–52. [aBH]
Viviani, P. & Terzulo, C. A. (1982) Trajectory determines movement dynamics. *Neuroscience* 7:431–37. [aBH]
Von Hofsten, C. (1985) Perception and action. In: *Goal-directed behavior: The concept of action in psychology,* ed. M. Frese & J. Sabini. Erlbaum. [aBH]
Von Holst, E. & Mittelstaedt, H. (1950) Das Reafferenzprinzip. *Naturwissenschaften* 37:464–76. [PC, aBH]
Von Weizsäcker, V. (1940) *Der Gestaltkreis.* Springer. [AFS]
Vu, K.-P. L. & Proctor, R. W. (2001) Determinants of right-left and top-bottom prevalence for two-dimensional spatial compatibility. *Journal of Experimental Psychology: Human Perception and Performance* 27:813–28. [RWP]
Walker, E. L. (1969) Reinforcement – "The one ring." In: *Reinforcement and behavior,* ed. J. T. Tapp. Academic Press. [rBH]
Wallace, R. A. (1971) S-R compatibility and the idea of a response code. *Journal of Experimental Psychology* 88:354–60. [aBH]
(1972) Spatial S-R compatibility effects involving kinesthetic cues. *Journal of Experimental Psychology* 93:163–68. [aBH]
Ward, L. M. (1982) Determinants of attention to local and global features of visual forms. *Journal of Experimental Psychology: Human Perception and Performance* 8:562–81. [aBH]
Warren, W. H. (1984) Perceiving affordances: Visual guidance of stair climbing. *Journal of Experimental Psychology: Human Perception and Performance* 10:683–703. [MJR]
(1998) Visually controlled locomotion: 40 years later. *Ecological Psychology* 10:177–220. [rBH]
Warrick, M. J. (1947) *Direction of movement in the use of control knobs to position visual indicators.* USAF AMC Report No.694-4C, Wright Airforce Base, Ohio, USA. [AFS]
Weimer, W. R. & Palermo, D. S., eds. (1974) *Cognition and the symbolic processes, vol. 1.* Erlbaum. [rBH]
(1982) *Cognition and the symbolic processes, vol. 2.* Erlbaum. [rBH]
Weiss, P. (1967) One plus one does not equal two. In: *The neurosciences,* ed. G. Querton, T. Melnechuk & F. Schmitt. Rockefeller University Press. [RES]
Welch, R. B. (1978) *Perceptual modification. Adapting to altered sensory environments.* Academic Press. [aBH]
(1986) Adaptation of space perception. In: *Handbook of perception and human performance, vol. 1: Sensory processes and perception,* ed. K. R. Boff, L. Kaufman & J. R. Thomas. Wiley. [aBH]
Welford, A. T. (1952) The "psychological refractory period" and the timing of high-speed performance – a review and a theory. *British Journal of Psychology* 43:2–19. [rBH]
(1968) *Fundamentals of skill.* Methuen. [aBH, RWP]
(1976) *Skilled performance: Perceptual and motor skills.* Scott-Foresman. [RWP]
(1980) Choice reaction time: Basic concepts. In: *Reaction times,* ed. A. T. Welford. Academic Press. [aBH]
Werner, H. & Wapner, S. (1952) Towards a general theory of perception. *Psychological Review* 59:324–38. [AFS]
Whiting, H. T. A., ed. (1984) *Human motor actions – Bernstein reassessed.* North-Holland/Elsevier. [rBH]
Williams, L. M. (1996) Cognitive inhibition and schizophrenic symptom subgroups. *Schizophrenia Bulletin* 22:139–51. [RDO]
Wise, S. P. & Murray, E. A. (2000) Arbitrary associations between antecedents and actions. *Trends in Neurosciences* 23:271–76. [PC]
Wolff, P. (1987) Perceptual learning by saccades: A cognitive approach. In: *Perspectives on perception and action,* ed. H. Heuer & A. F. Sanders. Erlbaum. [aBH]
(1999) Space perception and intended action. In: *Cognitive contributions to the perception of spatial and temporal events,* ed. G. Aschersleben, T. Bachmann & J. Müsseler. Elsevier. [aBH]
Wolpert, D. M. & Kawato, M. (1998) Multiple paired forward and inverse models for motor control. *Neural Networks* 11:1317–29. [DAR]
Wühr, P. & Müsseler, J. (2001) Time course of the blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human Perception and Performance* 27(5):1260–70. [JD, arBH, CO]
Yaniv, I., Meyer, D. E., Gordon, P. C., Huff, C. A. & Sevald, C. A. (1990) Vowel similarity, connectionist models, and syllable structure in motor programming of speech. *Journal of Memory and Language* 29:1–26. [aBH]
Zacks, J. M. & Tversky, B. (2001) Event structure in perception and conception. *Psychological Bulletin* 127:3–21. [rBH, JMZ]



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press

Zacks, J. M., Tversky, B. & Iyer, G. (2001) Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General* 130:29–58. [JMZ]

Ziessler, M. (1990) Task dependent formation of object features. *Studia Psychologica* 32:41–52. [GW]

Zinyuk, L., Kubik, S., Kaminsky, Y., Fenton, A. A. & Bures, J. (2000) Understanding hippocampal activity by using purposeful behavior: Place navigation induces place cell discharge in both task-relevant and task-irrelevant spatial reference frames. *Proceedings of the National Academy of Sciences USA* 97:3771–76. [JH]

Zwaan, R. A. & Radvansky, G. A. (1998) Situation models in language comprehension and memory. *Psychological Bulletin* 123:162–85. [JMZ]



https://doi.org/10.1017/S0140525X01330108 Published online by Cambridge University Press